DOI: https://doi.org/10.46632/jeae/1/2/3

Journal on Electronic and Automation Engineering
Vol: 1(2), 2022
REST Publisher; ISSN: 2583-6951
Website: http://restpublisher.com/journals/jeae/

Securing Wireless Sensor Networks Using Machine Learning

***D. Chitra, S. Ponlatha, M. Shanmugam, R. Kanagaraj**
Mahendra Engineering College, Namakkal, Tamil Nadu, India.
*Corresponding author email: chitrad@mahendra.info

Abstract: Wireless Sensor Networks (WSNs) are essential components of the Internet of Things (IoT), yet because of their limited resources and changeable network architectures, they are extremely susceptible to cyberattacks. The ineffectiveness of traditional security measures frequently makes the use of cutting-edge methods like Machine Learning (ML) necessary. This study offers a thorough examination of cybersecurity threats in WSNs, emphasising cutting-edge machine learning-based techniques for attack prevention, anomaly detection, and intrusion detection. The effectiveness of several machine learning methods, such as Random Forest (RF), Decision Tree (DT), Support Vector Machines (SVM), and Multi-Layer Perceptron, in identifying cyberthreats while minimising computing cost is assessed. To improve network resilience, a security framework with ML-driven detection and preventive procedures is suggested. Results from experiments show that machine learning approaches are more effective than traditional methods at accurately identifying security threats. The paper also emphasises the necessity for lightweight and optimised machine learning systems for securing WSNs, highlighting important design issues like as energy efficiency, computing limitations, and real-time adaptability.

Keywords: Wireless Sensor Networks, Machine Learning, Internet of Things, Denial-of-Service.

1. Introduction

The IoT is made possible by WSNs, which are inexpensive and low power [1]. Despite their vulnerability to attacks, these platforms have been studied for a variety of applications, including civilian, military, ecological, and healthcare [2]. In fact, there are serious security risks with WSNs. The use of energy-constrained devices, the wireless communicating channels that have been adopted, the use of multi-hop relays, the dynamic network topology, the varying medium-to-large network scales, the development of heterogeneous sensor nodes, and—above all—the various routing protocols that are used are all causes for concern. Developing preventive, identification, and mitigating safety measures for WSNs is crucial for creating safe and dependable IoT systems, as the latter includes a number of WSNs [3]. As a result, protecting WSNs is related to protecting IoT. Due to its portability, affordability, and simplicity of deployment, WSNs are among the best techniques for a variety of real-time applications [4]. The WSN is responsible for monitoring the area of interest, gathering data, and sending it to the point of access (base station) for analysis after post-processing. Some WSN systems make use of a high number of sensor nodes. Additionally, the memory and battery life of these wireless nodes are constrained. In order to maximise the benefits of these WSNs, a management system that can control the link between the WSN nodes and the point of access is necessary.

Since energy usage and security have a detrimental impact on one another, they are two of the most significant issues facing WSNs. Nodes in a WSN with more advanced security systems use more power, and vice versa. The requirement for both (lowering security and energy usage) is one of the issues that modern research in this area is tackling, given the difficult conditions under which these kinds of sensors can function [5,6]. The Triangle, which stands for Confidentiality, Integration, and Authentication (CIA), is another example of the traditional security procedures that need to be re-examined. Data encryption among two devices that communicate (two nodes) and related processes, like encryption and key exchange, are also regarded as conventional [7]. Furthermore, these approaches are energy-intensive, particularly given the frequent changes in network topologies brought on by the constant mobility of WSN nodes, as discussed in the preceding paragraph. Therefore, the goal is to find simpler and faster alternative technologies. Algorithms based on artificial intelligence, for instance, are among the techniques available for this use. Interaction with neighbouring WSN nodes, virus detection, packet analysis, node-to-node authentication, and availability maintenance are among abilities that a node can learn [8]. WSN

devices limit the network's power, storage spaces. computation and interaction capabilities, traditional WSN safety measures like spread spectrum, cryptography, and key management may not be effective at detecting attacks and may require complex software and hardware modifications, making them not enough to address WSN security issues [9]. Novel security paradigms have garnered increasing attention, and cybersecurity firms have invested up to USD 119 billion to address these issues [10]. This has resulted in newly developed methods to use ML to fortify WSN security against potential cyberattacks.

A sample of data, known as "training data," is used by algorithms to generate a mathematical representation in ML, one of the most well-known applications of artificial intelligence to date. This allows computers to make decisions or predictions without the need for explicit programming. The following factors make the ML type of WSNs significant: WSN ecosystems are so complicated that mathematical frameworks cannot be developed for them. Additionally, for some programs to function properly, data sets must be connected. Furthermore, ML techniques don't need human involvement, which is consistent with the nature of WSNs. Lastly, WSNs exhibit surprising dynamics and behaviours. ML in WSNs has two major obstacles: the necessity for big data sets for learning, and the mathematical and resource constraints of nodes. Regarding WSN network security, one of the biggest obstacles to ML algorithms is their inability to be applied to security needs that require secrecy and integrity. Consequently, ML methods can aid in enhancing wireless network security, mitigating various types of congestion issues [12], facilitating physical layer authentication procedures, and detecting errors. Additionally, ML algorithms are quite effective at identifying suspicious nodes and analysing packets as they move between WSN nodes.

It is crucial to implement a thorough and multifaceted strategy that makes use of cutting-edge ML methodologies in order to improve WSN security [13]. However, there are still a number of problems with ML integration into WSN. Furthermore, each WSN security architecture's statistical analysis is completely examined. As AI became more and more popular in the 1950s, ML swiftly rose to prominence and became a key component of the subject. Computer-executable algorithms for categorisation, sorting, regression, and optimisation are frequently employed to address a range of engineering, computing, and medical problems. These procedures have become simpler to execute as time has gone on and technological advancements have been made. As the algorithms converged, the frequency of this event increased. In today's world, ML is a very significant and intriguing technique. Some WSNs can carry out security-related tasks without requiring ML methods. An organisation or entity's capacity to run itself autonomously—that is, without outside intervention or control—is referred to as self-government, or autonomous administration. However, due to the low processing power and resources, conventional network security techniques are inadequate for adequately safeguarding WSNs [14,15]. It is commonly known that using ML approaches to improve WSN security is effective. The concept of "user authorisation" is considered unsuitable in the current situation.

2. Literature Review

Elsadig et al., have stated that WSNs are rapidly expanding in many different industries because of their special performance and features. WSNs are extremely susceptible to security threats, particularly Denial-of-Service (DoS) assaults. Finding WSN constraints, vulnerabilities, and security risks—with a particular emphasis on DoS attacks—is the main goal of this work. It examines current methods for identifying DoS assaults and suggests a lightweight ML strategy that combines the Gini feature identification technique with a DT algorithm. The method outperforms other classifiers in terms of accuracy, achieving 99.5% with a low processing overhead. [16]

According to Lai et al., WSNs have security issues that can deplete their meagre energy supplies. Communication and resource limitations make traditional security techniques insufficient. In order to identify DoS assaults in WSNs, this study suggests an online learning-based method. It presents an online aggressive or passive-a multi-class classifier that is noise-tolerant and a feature selection technique. The approach exhibits competitive performance when assessed using accuracy, precision, recall, and F1-score. [17]

Pandey et al., have suggested that the ability of ML to effectively identify network threats, particularly those in Internet of Things networks, has grown in popularity. However, because WSNs have limited computing and transmission power, conventional network intrusion detection techniques are not directly applicable to WSNs. At the moment, ML models for traffic data analysis are being intensively studied by researchers in the field of WSN intrusion detection. WSN networks produce high-dimensional traffic data as their user base and size grow, which makes it difficult for traditional ML models to extract features and detect them accurately. The particular requirements of this application environment may not be met by these restrictions. When compared to traditional

models, ML models can lessen the computational load and better understand the features of data traffic, which improves detection accuracy [18,19].

According to Ifzarne et al., WSNs have expanded quickly across a range of applications; nevertheless, because they are deployed in hostile and unmanaged environments, they provide security issues. This work focusses on applying online learning classifiers for intrusion detection in WSNs. An automated passive-aggression classifier as well as gain ratio feature selection are combined in the suggested model. The model's ability to detect a variety of attacks is demonstrated by experiments conducted on the WSN-DS dataset, which show a 96% detection rate, an accuracy of 86% for scheduling assaults, and 99% accuracy for regular traffic. [20]

Saleh et al., have described that WSNs are essential to cyber-physical systems, yet they are vulnerable to cyberattacks. This study uses ML methods to improve intrusion detection in WSNs, including Gaussian Nave Bayes (GNB) and Stochastic Gradient Descent (SGD) algorithms. Singular value decomposition and principal component analysis are used to lessen the computational demand on basic traffic information. In intrusion detection tasks, the suggested SG-IDS model outperforms state-of-the-art methods with an accuracy rate of 96% on the WSN-DS dataset. To further confirm its efficacy, it also performs exceptionally well in an evaluation of an Internet of Medical Things (IoMT) dataset. [21]

According to Annapurna et al., WSNs are frequently employed in monitoring applications; however, they are vulnerable to security risks, especially from malevolent nodes that launch security assaults. Unauthorised data access may result from these attacks, which take place on many network layers. Such threats include Sybil attacks and wormholes. The WHASA dataset is used to identify assaults using a prediction module that uses a variety of ML methods, including as XGBoost, Adaboost, Random Forest, and KNN, in order prevent additional harm. The objective is to categorise network access as either "standard" or "under attack" due to Sybil and Wormhole attacks. [22]

3. Research Methodology

To investigate cybersecurity risks in WSNs and evaluate the efficacy of ML based security measures, a methodical methodology was used. To investigate common attack types, current ML-driven intrusion detection and prevention methods. A security framework that emphasises computational efficiency and adaptability and incorporates ML techniques for cyberattack detection and prevention was developed using insights from the literature. The methodology comprised defining important design factors, assessing the performance of ML algorithms, and classifying them according to their use in WSN security. The capacity of several ML models to identify irregularities and stop attacks was evaluated. The viability of incorporating these methods into WSN environments with limited resources was examined. To improve overall system resilience, security issues, computing constraints, and optimisation techniques were investigated.

Types of Attacks in WSNs

Comparing WSNs to other sensory devices, they are very cheap and have limited processing resources. However, they are susceptible to a number of security risks due to their constrained processing capacity. WSNs are extremely vulnerable to physical attacks and unauthorised access because they are frequently installed in isolated, deserted areas. The following describes some of the most prevalent attack types in WSNs:

DoS (Denial of Service) Attack

These attacks interfere with network operations by overloading nodes with traffic or exhausting their resources, which stops authorised users from using the network.

Probe Attack

Probe attacks aim to get private data about the network, such as topology, vulnerabilities, or device configurations, that can be utilised for additional exploitation.

U2R (User to Root) Attack

This kind of attack involves a user-level attacker taking advantage of system flaws to take over a device or node as root or administrator. This permits essential data manipulation and unapproved network alterations.

R2L (Remote to Local) Attack

An external attacker gains unauthorised access to a network node in this attack. In order to obtain local access and take control of the victim's device, the hacker remotely takes advantage of vulnerabilities.

WSN Design Challenges and Unique Characteristics

The distinctive features of WSNs set them apart from conventional networks. These traits, along with limited resources, create a number of design issues that need to be resolved for effective functioning.

Unique Characteristics of WSNs

Resource-constrained sensor nodes with confined power, processing, and memory capacities make up WSNs. These networks function in dynamic settings where node movement or energy depletion regularly causes topological changes. Strong error-handling procedures are necessary because communication is frequently unreliable owing to interference and packet loss. WSNs are application-specific, self-organising, and energy-efficient because their nodes are hard to replace and run on batteries. Furthermore, open communication channels and unattended deployment make them vulnerable to attacks, which raise serious security concerns.

Design Challenges in WSNs

The main issue is energy management since extended network operation requires effective power consumption. Techniques for data aggregation ensure precise data delivery while minimising unnecessary transmissions. Managing dynamic topologies and resource limitations requires dependable communication and effective routing methods. While fault tolerance guarantees continuous operation even in the event of node failures, node placement choices have an impact on network coverage and connectivity. Synchronisation is necessary for time-sensitive applications, and security measures like encryption and authentication are critical to thwarting attackers. Developing effective WSNs requires addressing these issues, particularly for uses like military surveillance, industrial automation, and healthcare. Figure 1 shows the distinctive features of WSNs.

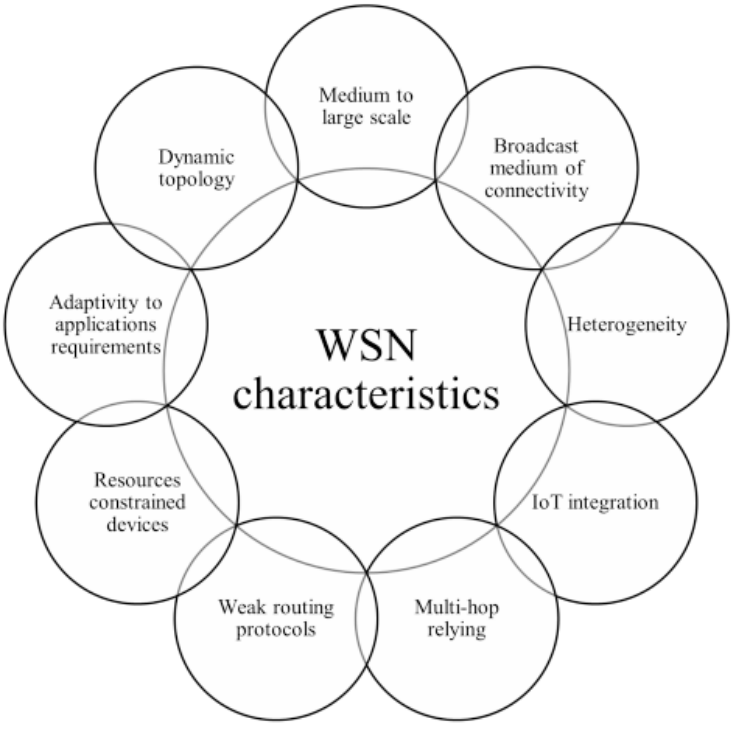

FIGURE 1. Distinctive features of WSNs

Cyber-attacks in WSNs

WSNs are extensively employed in vital applications like environmental monitoring, military surveillance, and healthcare; nevertheless, because of their wireless nature and limited resources, they are susceptible to a variety of cyberattacks. Common dangers include wormhole attacks, which generate false communication links to trick

routing protocols, Sybil attacks, which create many phoney identities to affect network operations, and Denial of Service (DoS) attacks, which exhaust node resources and interfere with communication. By intercepting and altering sent data, other attacks like sinkhole and selective forwarding jeopardise data integrity. The security and validity of sensitive data are further threatened by replay attacks and eavesdropping. Data loss, decreased network efficiency, and impaired application performance are all possible outcomes of these cyberthreats. Security mechanisms including intrusion detection systems, lightweight encryption methods, secure routing protocols, and trust-based frameworks are crucial for reducing these threats. Furthermore, anomaly detection based on ML can improve threat identification while using less energy. Reliability and resilience of WSN installations in practical applications depend on addressing these security issues.

ML Algorithms and its Types

An algorithm for ML is a computer technique that allows systems to recognise patterns in data and anticipate or decide without explicit programming. These algorithms examine input data, find underlying patterns, and train themselves to perform better over time. Supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning are the four primary categories into which ML algorithms are divided based on how they learn and handle data.

Supervised learning algorithms

It uses labelled data, which means that every input has a known output. By mapping inputs to the appropriate outputs, the model gains knowledge and improves its predictions in response to mistakes. Decision trees, which categorise data by dividing it into logical decisions, and linear regression, which predicts continuous values, are two examples. Neural networks are intended to replicate the human brain's capacity to identify intricate patterns, but Logistic Regression (LR) and SVM are frequently used in classification tasks.

Unsupervised learning algorithms

It is not dependent on data that has been labelled. Rather, they examine input data to find underlying structures, correlations, or patterns. These techniques are frequently applied to anomaly detection, dimensionality reduction, and clustering. For instance, Principal Component Analysis (PCA) minimises data dimensions while maintaining crucial information, whereas K-means clustering combines data points according to similarity. Neural networks that learn effective data representations are called autoencoders, and hierarchical clustering creates nested groupings.

Semi-supervised learning

This approach combines a small amount of labelled data with a larger set of unlabelled data to bridge the gap between supervised and unsupervised approaches. This method reduces the requirement for large labelled datasets while increasing learning efficiency. While graph-based algorithms use associations between data points to increase classification accuracy, self-training techniques iteratively label additional data based on an initial limited labelled sample.

Reinforcement learning

It is a distinct category where a device engages with its surroundings and gains knowledge by being rewarded or punished. Reinforcement learning, in contrast to other learning approaches, emphasises sequential decision-making in which past actions impact subsequent results. A model-free reinforcement learning technique called Q-learning aids an agent in gradually figuring out the optimal path of action. For increasingly complicated decision-making tasks, Deep Q Networks (DQN) integrate deep learning and reinforcement learning, whereas policy gradient approaches directly optimise policies to improve learning effectiveness. Figure 2 shows the flow diagram of ML Model.

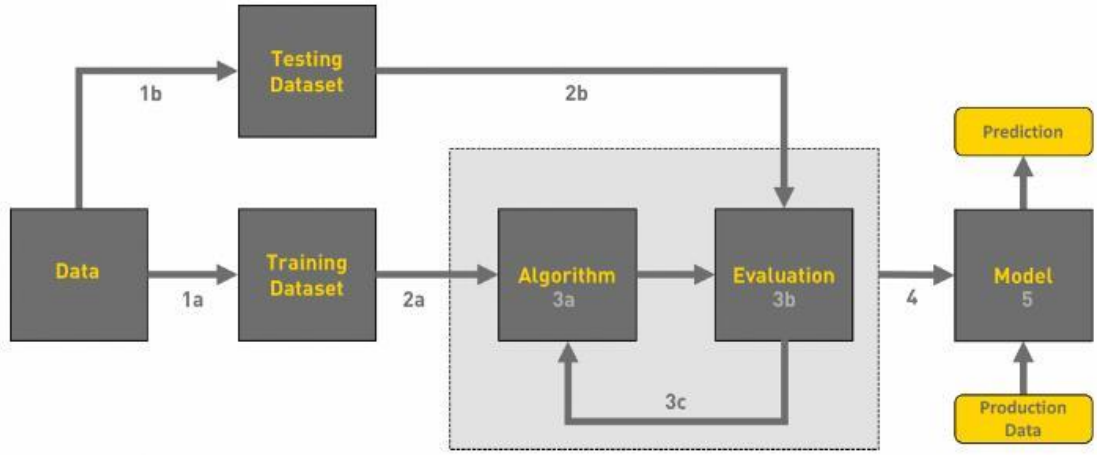

FIGURE 2. Flow diagram of ML Model

4. Result and Discussion

First, binary classification has been carried out, wherein typical (non-attack) samples are categorised into one group, and all attacks are categorised into a second group. Class 0 in binary class categorisation denotes normal, while class 1 denotes attacks. Then, a number of ML algorithms—such as DT, LR, RF, MLP, SVM, KNN, NB, and Voting Classifier—are used to detect WSN assaults. Precision, recall, F-1 score, weighted-average, and micro-average precision, ROC, and MCC are used to evaluate the performance of the aforementioned algorithms. The comparison graphs of several algorithms with varying performance metrics are displayed in Figure 3 & 4.

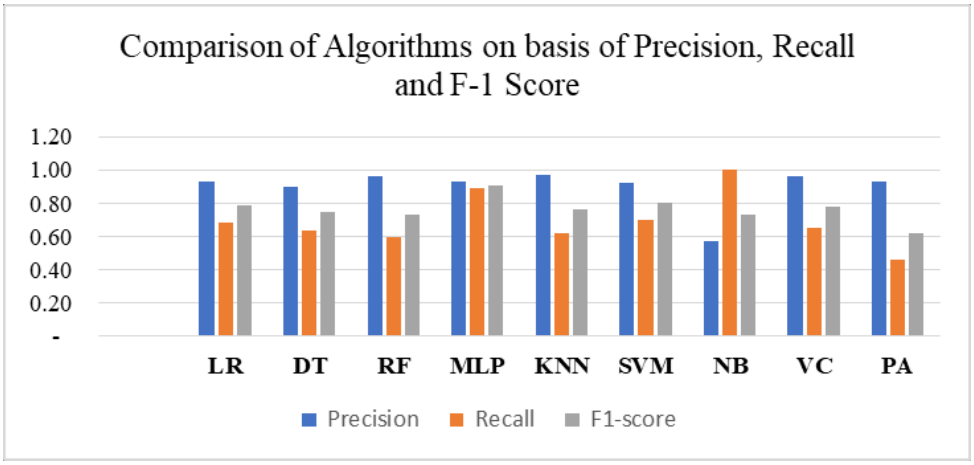

FIGURE 3. Comparison of various algorithms based on ML Classifiers

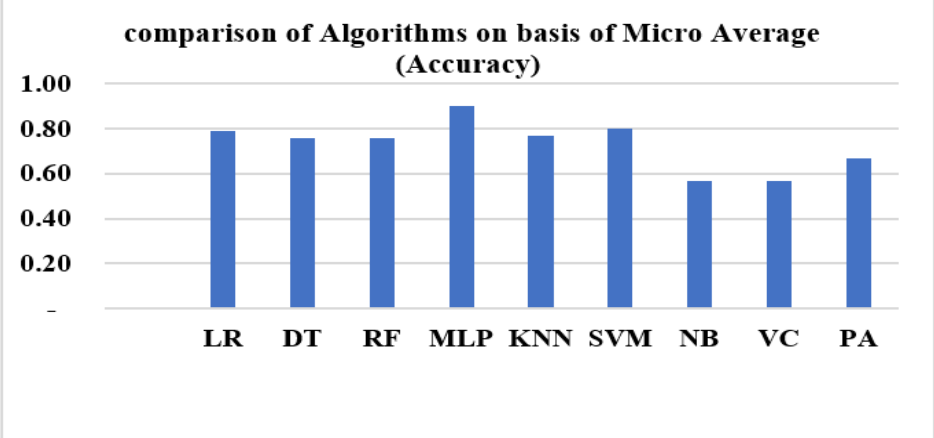

FIGURE 4. Accuracy-based comparison of algorithms

The two graphs use a number of evaluation metrics to compare how well different ML algorithms perform. Precision, recall, and F1-score—all of which are essential for comprehending an algorithm's capacity to accurately categorise positive occurrences while reducing false positives and false negatives—are compared in the first graph. In contrast to Naïve Bayes (NB), which catches the majority of positive examples at the expense of more false positives, MLP, RF, and SVM demonstrate great overall performance, attaining high precision and balanced recall. Micro-average accuracy, which gauges overall correctness across all classifications, is used in the second graph to assess the algorithms. Here, NB and VC perform comparatively poorly, while MLP attains the maximum accuracy, followed by LR, RF, and SVM. NB and VC, while having strong recall, have trouble with precision and overall accuracy, while MLP, RF, and SVM consistently show superior performance across several parameters, making them more dependable for classification.

5. Conclusion

ML makes it possible to detect intrusions, identify anomalies, and stop attacks, which improves the security of WSNs. While maintaining computational efficiency, the suggested ML-driven system incorporates detection and prevention techniques. According to experimental findings, ML algorithms like MLP, SVM, and RF outperform conventional techniques in terms of accuracy and precision. However, because WSNs have limited resources, striking a balance between security and energy efficiency is still difficult. It is crucial to design lightweight security measures and optimise ML models for real-time applications. While reducing computing overhead, adaptive learning techniques can improve detection accuracy. WSN defences against changing cyberthreats can be further strengthened by incorporating hybrid security techniques. For ML deployment to go well, network scalability and dynamic network issues must be resolved. WSN resilience across a range of applications can be enhanced by developments in ML-driven security frameworks. Enhancing energy-efficient ML models will contribute to strong security without sacrificing functionality. The long-term viability of safe WSNs will be aided by more research in this domain.

References

[1]. Abbas, G.; Mehmood, A.; Carsten, M.; Epiphaniou, G.; Lloret, J. (2022) Safety, Security and Privacy in Machine Learning Based Internet of Things. J. Sens. Actuator Netw. 11, 38.

[2]. Bajaj, K.; Sharma, B.; Singh, R. (2020) Integration of WSN with IoT applications: A vision, architecture, and future challenges. In Integration of WSN and IoT for Smart Cities; Springer: Berlin/Heidelberg, Germany; pp. 79–102.

[3]. Pavithran, D.; Shaalan, K.; Al-Karaki, J.N.; Gawanmeh, A. (2020) Towards building a blockchain framework for IoT. Clust. Comput. 23, 2089–2103.

[4]. Wazirali, R.; Ahmad, R.; Al-Amayreh, A.; Al-Madi, M.; Khalifeh, A. (2021) Secure Watermarking Schemes and Their Approaches in the IoT Technology: An Overview. Electronics 10, 1744.

[5]. Khashan, O.A.; Ahmad, R.; Khafajah, N.M. (2021) An automated lightweight encryption scheme for secure and energy-efficient communication in wireless sensor networks. Ad. Hoc. Netw. 115, 102448.

[6]. Ahmad, R.; Sundararajan, E.A.; Abu-Ain, T. (2021) Analysis the Effect of Clustering and Lightweight Encryption Approaches on WSNs Lifetime. In Proceedings of the 2021 International Conference on Electrical Engineering and Informatics (ICEEI), Kuala Terengganu, Malaysia, 12–13 October; IEEE: Selangor, Malaysia; pp. 1–6.

[7]. Rana, M.; Mamun, Q.; Islam, R. (2022) Lightweight cryptography in IoT networks: A survey. Futur. Gener. Comput. Syst. 129, 77–89.

[8]. Sharma, H.; Haque, A.; Blaabjerg, F. (2021) Machine learning in wireless sensor networks for smart cities: A survey. Electronics 10, 1012.

[9]. Hussain, F.; Hussain, R.; Hassan, S.A.; Hossain, E. (2020) Machine learning in IoT security: Current solutions and future challenges. IEEE Commun. Surv. Tutor. 22, 1686–1721.

[10]. Xu, L.D.; Lu, Y.; Li, L. (2021) Embedding Blockchain Technology Into IoT for Security: A Survey. IEEE Internet Things J. 8, 10452–10473.

[11]. Schwendemann, S.; Amjad, Z.; Sikora, A. (2021) A survey of machine-learning techniques for condition monitoring and predictive maintenance of bearings in grinding machines. Comput. Ind. 125, 103380.

[12]. Masdari, M. (2020) Energy Efficient Clustering and Congestion Control in WSNs with Mobile Sinks; Springer: Berlin/Heidelberg, Germany; Volume 111, ISBN 0123456789.

[13]. Schwendemann, S.; Amjad, Z.; Sikora, A. (2021) A survey of machine-learning techniques for condition monitoring and predictive maintenance of bearings in grinding machines. Comput. Ind. 125, Feb., Art. no. 103380.

[14]. Nayak, P.; Swetha, G.K.; Gupta, S.; Madhavi, K. (2021) Routing in wireless sensor networks using machine learning techniques: Challenges and opportunities. Measurement 178, Jun., Art. no. 108974.
[15]. Gouda, W.; Tahir, S.; Alanazi, S.; Almufareh, M.; Alwakid, G. (2022) Unsupervised outlier detection in IoT using deep VAE. Sensors 22, no. 17, p. 6617, Sep.
[16]. Elsadig, M.A. (2023) Detection of Denial-of-Service Attack in Wireless Sensor Networks: A lightweight Machine Learning Approach. IEEE Access.
[17]. Lai, T.T.; Tran, T.P.; Cho, J.; Yoo, M. (2023) DoS attack detection using online learning techniques in wireless sensor networks. Alexandria Engineering Journal 85, 307-319.
[18]. Pandey, J.K.; Kumar, S.; Lamin, M.; Gupta, S.; Dubey, R.K.; Sammy, F. (2022) A metaheuristic autoencoder deep learning model for intrusion detector system. Math. Probl. Eng. 2022, 3859155.
[19]. Almomani, O. (2021) A hybrid model using bio-inspired metaheuristic algorithms for network intrusion detection system. Comput. Mater. Contin. 68, 409–429.
[20]. Ifzarne, S.; Tabbaa, H.; Hafidi, I.; Lamghari, N. (2021) Anomaly detection using machine learning techniques in wireless sensor networks. In Journal of Physics: Conference Series, vol. 1743, no. 1, p. 012021. IOP Publishing.
[21]. Saleh, H.M.; Marouane, H.; Fakhfakh, A. (2024) Stochastic Gradient Descent Intrusions Detection for Wireless Sensor Network Attack Detection System Using Machine Learning. IEEE Access.
[22]. Annapurna, V.G.; Anusha, K.; Kamala Varsha, C.H.; Deepthi, M.; Keerthi, G. (2023) Detection of Network Layer Attacks in Wireless Sensor Network. Asian Journal For Convergence In Technology (AJCT) ISSN-2350-1146 9, no. 1, 42-48.

