Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. ICMASD (2025)

Intrusion Detection in Wireless Sensor Networks using SMOTE Tomek Link sampling technique

Gurpreet Kaur[1], Kamal Malik[1]

[1] Department of Computer Science & Engineering, CT University, Ludhiana

Article History:

Received: 12-12-2024

Revised: 25-01-2025

Accepted: 05-02-2025

Abstract: Wireless Sensor Networks (WSNs) are an essential component of cyber-physical systems, characterized by the integration of stationary and mobile sensors that collaboratively capture and transmit environmental data. These sensors utilize self-organization and multi-hop communication mechanisms, which enable them to efficiently gather and process information from their surroundings. Despite the advantages provided by WSNs, they are vulnerable to various attacks that can severely compromise their functionality, leading to a pressing need for effective intrusion detection systems. Traditional intrusion detection methods in WSNs face significant challenges, including low detection rates, imbalanced & complex higher dimensional datasets and false alarms. To address the aforementioned challenges, an innovative intrusion detection approach is proposed, which integrates Machine Learning (ML) techniques with the Synthetic Minority Oversampling Technique Tomek Link (SMOTETomekLink) algorithm. This hybrid methodology aims to synthesize minority instances within the dataset while removing Tomek links, thereby creating a balanced dataset that enhances the accuracy of intrusion detection systems tailored for WSNs. The application of feature scaling through standardization is another critical element of the proposed model. Furthermore, implementing the SMOTETomek resampling technique is essential for counteracting imbalances in WSN datasets, which addresses the persistent issues of overfitting and underfitting often observed in machine learning applications. A comprehensive evaluation of the proposed intrusion detection model was conducted using the Wireless Sensor Network Dataset (WSN-DS), which comprises 374,661 records. The findings achieved a remarkable accuracy rate of 99.78% in binary classification scenarios. This high level of accuracy demonstrates the efficacy and superiority of the developed system within the context of WSN intrusion detection.

Keywords: Intrusion Detection, Machine Learning, Wireless Sensor Network, SMOTE

1 Introduction

Wireless Sensor Networks (WSNs) are indeed a transformative technology that facilitates data collection, processing, and transmission from distributed sensor nodes. These nodes, equipped with various sensors and communication capabilities, play a critical role in monitoring and sensing environmental conditions. [1]. WSN security is of utmost importance due to the sensitive nature of the data transmitted and the potential vulnerabilities in the network [2]. WSNs face security threats, like unauthorized access and denial of service attacks [3]. The distributed and wireless nature of WSNs makes them more susceptible to these threats. Ensuring the confidentiality, integrity, and availability of data within WSNs is crucial to maintaining the trust and reliability of these networks [4]. Intrusion detection is a critical component of WSN security, aimed at detecting and mitigating malicious activities within the

network [5]. Rule-based intrusion detection systems (IDS) are increasingly viewed as inadequate for the demands of modern cybersecurity due to their reliance on predefined signatures and static thresholds. This reliance limits their capability in identifying sophisticated or novel attacks, particularly those that evolve rapidly. [6]. Machine Learning (ML) techniques have emerged as a promising approach for WSN intrusion detection [1]. ML algorithms, such as decision trees, random forests and gradient boosting methods, can extract valuable insights from complex WSN datasets, improving the accuracy and effectiveness of intrusion detection mechanisms [1, 7]. Existing works in WSN intrusion detection face many challenges. Limited scalability is a significant concern, as WSNs often involve a large number of sensor nodes, leading to increased computational complexity [8]. High false positive rates, imbalanced datasets and inadequate adaptability to evolving attack techniques are also challenges that need to be addressed [9–11]. Our proposed model, combines Machine Learning (ML) techniques with the Synthetic Minority Oversampling Technique Tomek (SMOTE-Tomek), offers a significant advancement in intrusion detection for WSNs (WSNs) [9, 12, 13]. This model holds great importance in addressing the limitations of existing approaches and providing an optimal solution for intrusion detection in WSNs. The primary motivation behind our work is to deal with imbalanced dataset in WSNs. This imbalance can compromise the accuracy of detection results. Our model aims to overcome this challenge by employing the SMOTE-Tomek technique, synthesizing minority instances and removing Tomek links to achieve a balanced dataset [14]. Additionally, our model uses the power of various ML algorithms, including Decision Trees (DT), Random Forests (RF) and K-Nearest Neighbor (KNN). These algorithms enable the development of a robust intrusion detection system capable of learning complex patterns and anomalies from WSN data. The remaining sections of this paper are structured as follows: In Section 2, we delve into a comprehensive review of the existing literature, with a specific focus on intrusion detection on WSNs. Section 3 discusses about research methodology, including an extensive description of the dataset employed. The experimental setup and performance evaluation are elaborated in Section 4. Finally, Section 5 encapsulates the conclusion and further enhancements for future research.

2 Related Works

In the work done by [9] introduced an approach that employs the Synthetic Minority Oversampling Technique (SMOTE) to address dataset imbalance, followed by training a classifier using the Random Forest algorithm to detect the intrusions. Simulations were performed on a standard intrusion dataset, demonstrating that the RF achieved an accuracy of 92.39%, surpassing other algorithms in comparison. Furthermore, by applying SMOTE to oversample the minority samples, the accuracy of the RF classifier improved to 92.57%. This enhancement highlights the effectiveness of SMOTE in mitigating the adverse effects of class imbalance, enabling the model to better recognize and classify minority samples within the dataset.

The work proposed in [10] employed intrusion detection model tailored to WSNs characteristics, using information gain ratio (IGR) and the online Passive Aggressive (PA) classifier. Experiments on a WSN-DS dataset resulted in the proposed model achieving a 96% detection rate for normal behavior or attacks. Detection accuracies were 86% for scheduling, 68% for Grayhole, 63% for Flooding, and 46% for Blackhole attacks, with 99% accuracy for normal traffic.

In [15], presented a data mining approach to detect various types of DoS attacks, where they applied several classification algorithms, including KNN, Naive Bayes, Logistic Regression, Support Vector Machine (SVM), and Artificial Neural Network (ANN), to the dataset and assessed their performance in identifying these attacks. The analysis revealed that ANN achieved the highest accuracy at 98.56%, followed closely by KNN at 98.4%. These results suggest that ANN and KNN are strong candidates for intrusion detection, making them suitable recommendations for network specialists and analysts.

[11] proposed a multitier intrusion detection framework for WSNs, implementing a security approach with two detection layers. The first layer, situated at the edge of the distributed network, utilizes a Naive Bayes classifier for quick decisions concerning inspected packets. The second layer, positioned in the cloud, employed a Random Forest multi-class classifier for thorough packet analysis. It achieved excellent performance scores, including a 100% precision rate for Normal attacks, 90.4% for Flooding attacks, 99.5% for Scheduling attacks, 97% for Grayhole attacks, and 99.9% for Blackhole attacks.

Work done by [16] introduced a deep learning model that utilized a convolutional neural network. This model encompassed two essential phases: detecting intrusions and preventing them. It acquired meaningful feature representations from a labeled dataset and performed accurate classifications. They harnessed the power of the convolutional neural network to prevent intrusions in WSNs. They employed the WSN-DS to assess the system's effectiveness. The test outcomes showed that the proposed system achieved a remarkable accuracy rate of 97%.

An optimized collaborative intrusion detection system (OCIDS) was developed by [17] for WSNs using an enhanced artificial bee colony optimization (BCO) algorithm. It improved the accuracy of intrusion detection and resource efficiency. The OCIDS also enhanced the weighted support vector machine (SVM) algorithm to reduce false alarms. The outcome of OCIDS outperformed other systems with a 97.9% detection rate and a 1.8% false alarm rate, demonstrating a clear advantage of using it.

In [18], it employed the use of XGBoost in Intrusion Detection Systems (IDS) for imbalanced data in WSNs cyber-attacks. To assess its performance, comparison of XGBoost to decision trees and naive Bayes was done, employing various evaluation metrics. The results showed XGBoost's superiority, with the highest AUC values across scheduling, normal, grayhole, flooding, and blackhole classes, achieving 98.7%, 99.63%, 99.94%, 99.97%, and 99.99%, respectively.

The work done by [19] proposed an autoencoder-based framework using convolutional and recurrent networks for cyber threat detection in IIoT networks. Fully connected networks leverage extracted features for attack event classification. Empirical results highlight the framework's efficacy, outperforming contemporary methods and showcasing suitability for real-world IIoT networks.

3 Methodology

Detailed description of the experimental procedures and implementation specifics are as follows:

• **Data Collection**: Raw data from WSNs was collected to form the foundation for the experimental dataset.

• **Preprocessing Techniques**: A crucial aspect of the methodology involved the preprocessing of the collected data. Standardization was applied to normalize input features. Further, label encoding was

done to convert categorical target feature into a numerical format, facilitating compatibility with machine algorithms.

• **Data Balancing**: To address the challenge of imbalanced datasets and potential overfitting, the Synthetic Minority Over-sampling Technique (SMOTE) combined with the Tomek links removal method (SMOTETomek) was utilized. This technique ensured a balanced representation of both normal and intrusion instances in the dataset, contributing to improved model generalization.

• **Data Splitting**: The experimental dataset underwent k-fold cross-validation, specifically with 10 folds, to split it into training and testing sets. This process enhances the model's performance.

• **Model Building**: Several machine learning algorithms like RF, DT and KNN were implemented for intrusion detection model development.

• **Model Evaluation:** Various evaluation metrics, including accuracy, precision, recall, and F1-score, were employed to showcase the effectiveness in detecting and classifying intrusions.

• **Model Selection**: The final step involved selecting the best-performing model based on the comprehensive evaluation results.

The overall algorithmic steps of our proposal for intrusion detection in WSNs is shown in Algorithm 1 as follows:

Algorithm 1: Procedure for IDS-WSN

```
procedure INTRUSIONDETECTIONWSN(file_path, target_column)
    Input: WSN-DS as a CSV file path (file_path)
    Output: Trained ML models, Evaluate Performance of each model
procedure READCSV (file_path)
    Initialize DataFrame:df
    Df ← Read CSV file located at file_path
 end procedure
procedure PREPROCESSDATA(df, target_coulmn)
    Initialize Features: X
    Initialize Target_Variable : y
    X,y ← extract features and target variable from df
  APPLYSTANDARDIZATION(X)
   X ← Standardize Numerical features in X
   ENCODELABELS(y)
```

```
        y ← Encode categorical labels in y using LabelEncoder
    end procedure
    procedure DATABALANCING (X,y)
        BALANCEDATA(X,y)
        X,y ← Apply SMOTETomek to balance the dataset
    end procedure
    procedure SPLITDATA(X,y)
        Initialize K-Fold Cross-Validator : k f
        k f ← Initialize KFold(n_splits = 10, shuffle = True)
          for train_index, test_index in k f. split (X,y) do
           X_train, X_test ←Split X into training and testing sets
           Y_train, y_test ← Split y into training and testing sets
           BUILDMODELS(X_train, y_train)
           Initialize each ML Model : DT, RF, KNN
           EVALUATEPERFORMANCE (X_test, y_test)
           Evaluate Performances for each ML Model
          end for
      end procedure
    end procedure
```

3.1 Dataset Description

The Wireless Sensor Network DoS Detection Dataset (WSN-DS) [20] is a comprehensive collection crafted for the specific purpose of identifying Denial-of-Service (DoS) attacks in WSNs (WSNs). This dataset, consisting of 374,661 records, was assembled using the LEACH protocol, a widely adopted routing protocol in WSNs, and includes instances of four types of DoS attacks (Blackhole, Grayhole, Flooding, and Scheduling) as well as normal network behavior scenarios.

3.2 Data Preprocessing

In this section, we describe the key data preprocessing steps we have applied to our dataset.

1. Standardization: It is the process of transforming data such that it has a mean of 0 and a standard deviation of 1. This step is essential when dealing with features that have different scales or units. In WSN intrusion detection, sensor data may have varying measurement units and scales. Standardization helps in bringing all features to a common scale, which is important for machine learning algorithms that rely on distance metrics or gradient-based optimization.

The formula for standardization is as follows:

$$X_{standardized} = \frac{X - \mu}{\sigma}$$

Where: • $X_{standardized}$ is the standardized value of feature X.

• X is the original feature value.

• μ is the mean of the feature X.

• σ is the standard deviation of the feature X.

2. Label Encoding: Label encoding is a technique used to transform categorical data into numerical labels. • Binary Classification For binary classification, where we have two classes, ”Normal” and ”Attack,” we perform label encoding as shown in Table 1:

Table 1. Label Encoding for Binary Classification

Class	**Label Encoding**
Normal	0
Attack	1

3.3 Data Balancing using SMOTE-TomekLink

This approach combines SMOTE and Tomek-Links oversampling and undersampling methods. SMOTE is an oversampling technique and a statistical method for balancing the class distribution by generating new synthetic minority class samples. This approach helps mitigate overfitting issues associated with random oversampling methods and has been widely adopted to address class imbalance problems. On the other hand, Tomek-Links is an undersampling technique designed to remove instances on the ”Tomek link,” which are pairs of data points from different classes that are close to each other in the dataset. Removing these pairs helps separate minority and majority classes, reducing noise in the majority class. By combining the SMOTE and Tomek-Links (STL), the study aims to effectively tackle the imbalanced class problem in WSN data, providing a balanced dataset for more accurate intrusion detection. This strategy enhances class separation, stabilizes data distribution, and ultimately improves the performance of intrusion detection systems in the context of WSNs. Now, in this case of binary classification, distribution of classes without and with SMOTE-Tomek-Link is shown in the table below:

Table 2. Binary Classification class distribution

Class	WoSTL	WiSTL
Normal	340,066	340,056
Attack	34,595	339,610

3.4 Machine Learning Algorithms used in the proposed approach

- Decision Trees (DT): Decision trees are versatile tools widely applied in various domains, including machine learning, image processing, and pattern recognition. A decision tree comprises essential components: the root node, branches, and leaf nodes. The root node represents the entire dataset, which is partitioned into homogenous subsets. Branches represent combinations of attributes, while leaf nodes mark the end of the decision-making process [21].
- Random Forest (RF): Random Forest, a meta-approximation technique, enhances accuracy through averaging. It prevents overfitting by fitting multiple decision tree classifiers to various subsets of the dataset. Each subset is chosen independently from the feature space, resulting in a set of uncorrelated Decision Trees derived from different training data points. Each tree predicts a class, and the majority vote among the trees determines the model's prediction [22].
- K-Nearest Neighbour (KNN): K-Nearest Neighbour is a simple yet effective supervised machine learning algorithm for classification and regression tasks. It operates on the principle of proximity, classifying data points based on the majority class among their nearest neighbors. KNN's flexibility allows it to adapt to various data distributions, making it a valuable choice for both simple and complex datasets.

4 Results and Discussion

4.1 Experimental Setup

The experiments are carried out on a system running Microsoft Windows 10, which is equipped with an Intel(R) Core(TM) i5 operating at 1.90GHz, complemented by a 500GB SSD and 8GB of RAM. The implementation of the proposed model was done using the Python programming language, making use of a selection of commonly utilized libraries, including Pandas, NumPy, Matplotlib, Seaborn, TensorFlow, Keras, Scikit-learn, and others.

4.2 Performance Evaluation Metrics

The evaluation of our proposed model involved the utilization of multiple performance metrics to assess its effectiveness. These metrics are defined as follows:

- Confusion Matrix: Table 3 shows the confusion matrix where TP represents True Positive, TN represents True Negative, FP stands for False Positive, and FN denotes False Negative.
- Accuracy: It is defined as

$$\text{Accuracy} = \frac{TP + TN}{T P + F P + F N + T N} \quad (1)$$

- Precision: It is defined as

$$\text{Precision} = \frac{TP}{T P + F P} \quad (2)$$

- Recall: It is defined as

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

- F1-Score: It is defined as

$$\text{F1-score} = 2 * \frac{\text{Precision} * \text{Recall}}{}$$

Precision + Recall (4)

- ROC Curves are two-dimensional plots commonly employed for evaluating classifier effectiveness. An AUC (Area under the Curve) value approaching 1 indicates strong class separability, while an AUC value approaching 0 indicates suboptimal performance.

4.3 Result Analysis

In the binary results analysis, we evaluated the performance of different ML algorithms for intrusion detection in WSNs. Two experiments were conducted: one with SMOTETomek-Link (WiSTL) and another without SMOTETomek-Link (WoSTL). In evaluating the binary performance of accuracy, precision, recall, and f1- score performances of various ML models for intrusion detection in WSNs, the presented table below reveals performance across all techniques. In the scenario of WoSTL, DT, RF and KNN achieved accuracies of 99.52%, 99.69% and 99.63% respectively. The corresponding precision values of 98.63%, 99.26%, and 99.17% respectively. The corresponding recall values for these models are 98.57%, 98.9%, and 98.68% respectively. The corresponding F1-score values of 98.6%, 99.08%, and 98.92% respectively. When applying data balancing with SMOTETomek (WiSTL), accuracy, precision, recall, and f1-score values generally improved. In the scenario of WiSTL, DT, RF, and KNN achieved accuracies of 99.65%, 99.78%, and 99.5% respectively. The corresponding precision values of 99.65%, 99.78%, and 99.5% respectively. The corresponding recall values for these models are 99.65%, 99.78%, and 99.5% respectively. The corresponding F1-score values of 99.63%, 99.74%, and 99.5% respectively. Among these, Random Forest consistently demonstrated the highest performance scores in all scenarios, making it the model providing superior performance for intrusion detection in WSNs in the given context.

Table 3. Binary performance analysis of with and without data balancing using SMOTETomek

Technique	ML	Accuracy	Precision	Recall	F1-score
WoSTL	DT	99.52	98.63	98.57	98.6
	RF	99.69	99.26	98.9	99.08
	KNN	99.63	99.17	98.68	98.92
WiSTL	DT	99.65	99.65	99.65	99.63
	RF	99.78	99.78	99.78	99.78
	KNN	99.5	99.5	99.5	99.5

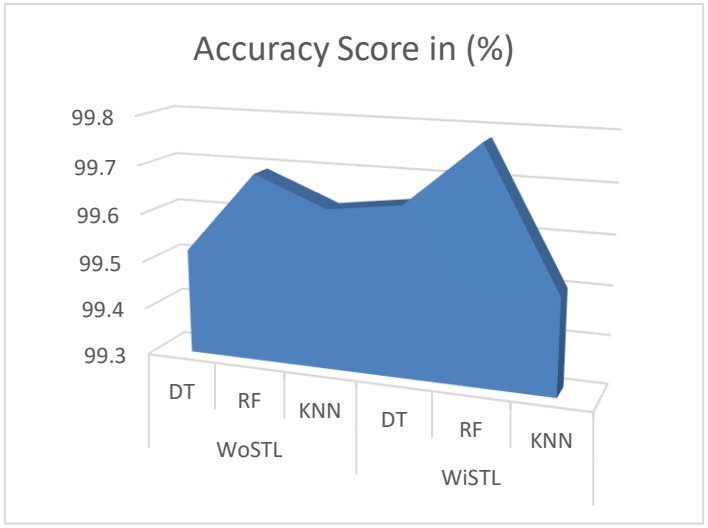


Fig. 1. Binary accuracy analysis of ML model for wireless sensor networks

Figure 1 compares the accuracy in graphical form of various ML models on WoSTL and WiSTL models. However, in WiSTL, where the SMOTETomek-Link technique was employed for data balancing, the accuracy rates of ML algorithms were improved. From Figure 2, it is evident that RF outperforms other algorithms in terms of performance metrics in the WiSTL experiment for intrusion detection in WSNs. RF consistently achieves the highest accuracy rate of 99.78%, indicating its ability to correctly classify instances as either normal or intrusions.

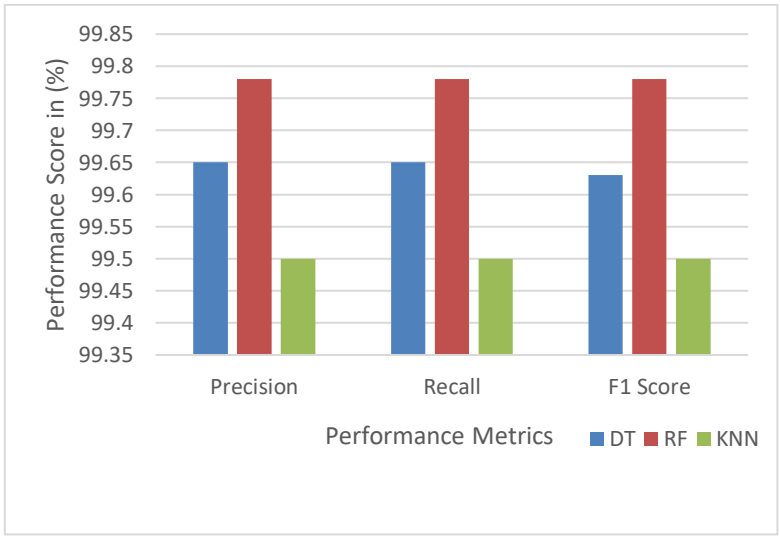
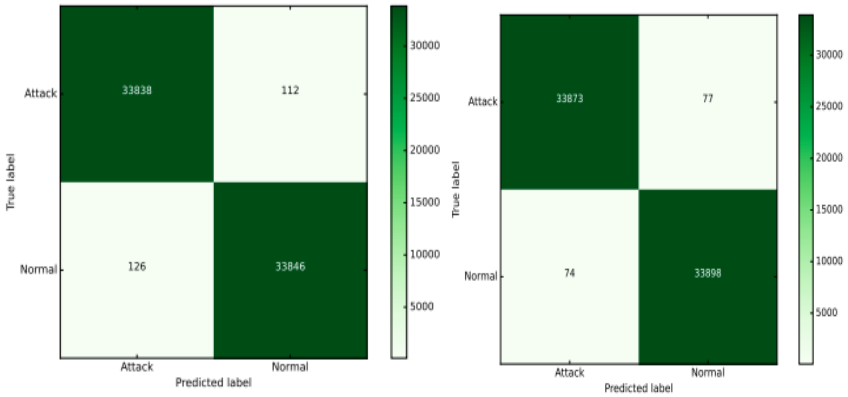
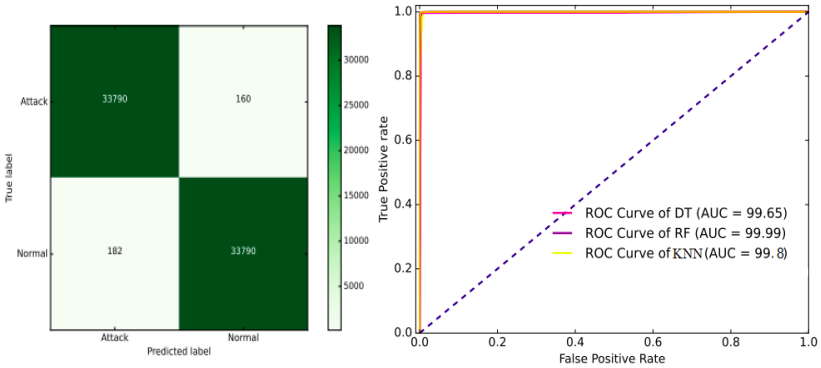


Fig. 2. Binary Performance Analysis for WSN

From Figure 3, RF demonstrates higher true positive and true negative rates compared to other algorithms, indicating its superior ability to correctly identify both normal instances and intrusions. Specifically, RF achieves a true positive rate of 33,873 and a true negative rate of 33,898, suggesting its effectiveness in accurately detecting intrusions. In contrast, RF exhibits lower false positive and false negative rates, with only 77 false positive and 74 false negative instances.



a) DT b) RF



c) KNN

Fig.3 Confusion Matrix **Fig. 4 ROC curve for binary classification**

Furthermore, the ROC curve in Figure 4 demonstrates the AUC score, which serves as a measure of the overall performance of the ML algorithms. RF achieves an impressive AUC score of 99.99%, by proving its superiority in detecting intrusions. The high AUC score indicates that RF exhibits a high true positive rate while maintaining a low false positive rate, making it an ideal choice for intrusion detection in WSNs.

5 Conclusion

In conclusion, proposed intrusion detection model, combining machine learning techniques with the SMOTE-TomekLink resampling method effectively balanced imbalanced WSN dataset. Feature scaling through standardization is employed to normalize input features, enhancing the model's accuracy and robustness. The result of findings is the remarkable performance achieved by our model, with an accuracy rate of 99.78% in binary classification. These results highlight the effectiveness and superiority of proposed model. Future research in WSN intrusion detection can explore hybrid feature selection to reduce computational complexity, employ deep learning models, especially fine-tuned models in IDS, and consider a hierarchical approach in WSNs to improve performance.

References

[1] Gebremariam, G.G., Panda, J., Indu, S.: Design of advanced intrusion detection systems based on hybrid machine learning techniques in hierarchically wireless sensor networks. Connection Science 35(1), 2246703 (2023).

[2] Yakubu, M.M., Maiwada, U.D.: Resource limitations for wireless sensor networks to establish a comprehensive security system in the 5g network. UMYU Scientifica 2(2), 44–52 (2023).

[3] Nimbalkar, A.D., Azmat, A., Patil, Y.: Security issues in wireless sensor networks. i-Manager's Journal on Wireless Communication Networks 11(2), 32 (2023)

[4] Alghamdi, R., Bellaiche, M.: A cascaded federated deep learning based framework for detecting wormhole attacks in iot networks. Computers & Security 125, 103014 (2023)

[5] Heidari, A., Jabraeil Jamali, M.A.: Internet of things intrusion detection systems: A comprehensive review and future directions. Cluster Computing, 1–28 (2022)

[6] Sezgin, A., Boyacı, A.: Aid4i: An intrusion detection framework for industrial internet of things using automated machine learning. Computers, Materials & Continua 76(2) (2023).

[7] Talukder, M.A., Hasan, K.F., Islam, M.M., Uddin, M.A., Akhter, A., Yousuf, M.A., Alharbi, F., Moni, M.A.: A dependable hybrid machine learning model for network intrusion detection. Journal of Information Security and Applications 72, 103405 (2023).

[8] Sharmin, S., Ahmedy, I., Md Noor, R.: An energy-efficient data aggregation clustering algorithm for wireless sensor networks using hybrid pso. Energies 16(5), 2487 (2023).

[9] Tan, X., Su, S., Huang, Z., Guo, X., Zuo, Z., Sun, X., Li, L.: Wireless sensor networks intrusion detection based on smote and the random forest algorithm. Sensors 19(1), 203 (2019)

[10] Ifzarne, S., Tabbaa, H., Hafidi, I., Lamghari, N.: Anomaly detection using machine learning techniques in wireless sensor networks. In: Journal of Physics: Conference Series, vol. 1743, p. 012021 (2021). IOP Publishing

[11] Alruhaily, N.M., Ibrahim, D.M.: A multi-layer machine learning-based intrusion detection system for wireless sensor networks. International Journal of Advanced Computer Science and Applications 12(4), 281–288 (2021).

[12] Zhang, H., Huang, L., Wu, C.Q., Li, Z.: An effective convolutional neural network based on smote and gaussian mixture model for intrusion detection in imbalanced dataset. Computer Networks 177, 107315 (2020)

[13] Mohammadi, S., Babagoli, M.: A novel hybrid hunger games algorithm for intrusion detection systems based on nonlinear regression modeling. International Journal of Information Security, 1–19 (2023)

[14] Chandra, W., Suprihatin, B., Resti, Y.: Median-knn regressor-smote-tomek links for handling missing and imbalanced data in air quality prediction. Symmetry 15(4), 887 (2023).

[15] Rezvi, M.A., Moontaha, S., Trisha, K.A., Cynthia, S.T., Ripon, S.: Data mining approach to analyzing intrusion detection of wireless sensor network. Indonesian J. Electric. Eng. Comput. Sci 21(1), 516–523 (2021).

[16] Chandre, P., Mahalle, P., Shinde, G.: Intrusion prevention system using convolutional neural network for wireless sensor network. Int J Artif Intell ISSN 2252(8938), 8938 (2022).

[17] Elsaid, S.A., Albatati, N.S.: An optimized collaborative intrusion detection system for wireless sensor networks. Soft Computing 24(16), 12553– 12567 (2020).

[18] Putrada, A.G., Alamsyah, N., Pane, S.F., Fauzan, M.N.: Xgboost for ids on wsn cyber attacks with imbalanced data. In: 2022 International Symposium on Electronics and Smart Devices (ISESD), pp. 1–7 (2022). IEEE.

[19] Khan, I.A., Moustafa, N., Pi, D., Sallam, K.M., Zomaya, A.Y., Li, B.: A new explainable deep learning framework for cyber threat discovery in industrial iot networks. IEEE Internet of Things Journal 9(13), 11604– 11613 (2021).

[20] Almomani, I., Al-Kasasbeh, B., Al-Akhras, M., et al.: Wsn-ds: A dataset for intrusion detection systems in wireless sensor networks. Journal of Sensors 2016 (2016).

[21] Ahmed, N., Ahammed, R., Islam, M.M., Uddin, M.A., Akhter, A., Talukder, M.A., Paul, B.K.: Machine learning based diabetes prediction and development of smart web application. International Journal of Cognitive Computing in Engineering 2, 229–241 (2021).

[22] Ahmad, M., Riaz, Q., Zeeshan, M., Tahir, H., Haider, S.A., Khan, M.S.: Intrusion detection in internet of things using supervised machine learning based on application and transport layer features using unsw-nb15 data-set. EURASIP Journal on Wireless Communications and Networking 2021(1), 1–23 (2021).