

Detecting network security incidents in wireless sensor networks using machine learning

Tamara Zhukabayeva^{1,2}, Atdhe Buja^{1,3}, Melinda Pacolli⁴, Yerik Mardenov^{1,5}

¹International Science Complex “Astana”, Astana, Kazakhstan

²Eurasian National University, Astana, Kazakhstan

³ICT Academy Research, Prishtina, Kosovo

⁴ECPD, Prishtina, Kosovo

⁵Department of Information Technology and Engineering, Astana International University, Astana, Kazakhstan

Article Info

Article history:

Received Apr 24, 2024

Revised Sep 9, 2024

Accepted Oct 7, 2024

Keywords:

Anomaly detection

Artificial intelligence

Cybersecurity

Internet of things

Wireless sensor networks

ABSTRACT

This study enhances the domain of cybersecurity within wireless sensor networks (WSNs) through the integration of sophisticated artificial intelligence (AI) and machine learning (ML) techniques. By conducting an exploratory data analysis (EDA), this research reveals critical insights into network behavior, facilitating the development of predictive models for anomaly detection. The application of ML algorithms decision trees (DT) and random forest (RF) demonstrated dominant performance in identifying potential security threats, as evidenced by metrics accuracy, precision, recall, and F1 scores. This work not only enhances the security framework for WSNs but also contributes to the extensive field of network security, offering a robust analytical and predictive methodology for future cybersecurity initiatives. The advanced model can be deployed in other WSN and internet of things (IoT) based applications.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Atdhe Buja

International Science Complex “Astana”

Astana, Kazakhstan

Email: atdhe.buja@academyict.net

1. INTRODUCTION

Wireless sensor networks (WSNs) have appeared as a key technology in varied domains, from environmental monitoring to industrial automation, outstanding to their ability to collect data from remote and tough environments. However, the broad deployment of WSNs has introduced new challenges, in terms of security, due to the vulnerabilities of wireless communication and resource-constrained IoT sensor devices. The utilization of machine learning (ML) and the internet of things (IoT) has engaged the focus of researchers following it can process vast amounts of data capable of unraveling various industry challenges [1]. Outcome, assuring the integrity, confidentiality, and availability of data transmitted within WSNs has become a concern for researchers and cybersecurity practitioners as well [2].

Several works have focused on approaches that have been proposed to detect and mitigate security threats in WSNs, ranging from cryptographic protocols to anomaly detection techniques. Through these, ML-based anomaly detection has gathered significant attention due to its capability to suit evolving threats and detect previously invisible attacks. Therefore, there is a need for a breakthrough force of AI-based models to mitigate cyber threats within IoT systems [3]. For instance, research work has investigated a secured framework to detect and stop data integrity attacks in WSNs in microgrids [4]. The breakout of the Mirai botnet uses IoT vulnerabilities and ruins numerous websites and domain name systems highlighting the need for a robust cybersecurity mechanism [5].

ML has been applied for anomaly detection systems in IoT systems and has been better [6]. Even IDS-based solutions present limitations in false positives [7]. Such research work provides a Hadoop-based framework to identify the malicious IoT traffic using a modified Tomek-link under-sampling integrated with automated Hyper-parameter tuning of machine learning classifiers [8]. So anomaly detection systems depend on a centralized management method to collect and process data generated by IoT devices, and a federated learning (FL) - base exhibits better accuracy [9]. Outlier detection is an important issue in IoT; in this study, a review of the state of the art in outlier detection approaches based on machine learning is given [10]. This research aims for goals:

- To develop a robust ML-based framework for detecting network security incidents in WSNs, using features such as communication metrics, environmental data, and energy consumption.
- Evaluate the performance of the developed framework, to assess its effectiveness in accurately identifying various types of network attacks.
- Identify appropriate ML methods to effectively detect network security incidents in WSNs.

This research introduces an innovative ML-based framework specially designed to detect network security incidents in WSNs. Distinct earliest approaches that focused on cryptographic solutions or rule-based anomaly detection methods, our framework integrates thorough feature engineering, including communication metrics, environmental data, and energy consumption, to advance the detection capabilities. The results of this study hand robust evidence supporting the application of ML techniques for advancing the security of WSNs. This solution could be used by network administrators and security analysts to accurately detect varied network security incidents in WSNs.

The following sections will illustrate the relevance and meaning of our work. The paper is structured as follows: section 2 presents materials and methods used during our experimentation to gain the results, including the data preprocessing steps, feature engineering techniques, and the ML models implemented. Section 3 discusses the experimentation and results, exhibiting the performance metrics of the decision tree (DT) and random forest (RF) models. This section will also include a comparative analysis to emphasize the accuracy and robustness of our proposed framework. In the end, section 4, will discuss the inferences of our findings, results, and future research directions, highlighting the practical applications and potential enhancements in the field of WSN security.

2. METHOD

2.1. Study area

The research aims to enhance cybersecurity within WSNs. WSNs are essential in many applications, from consumer appliances to industrial systems. By their critical purpose, assuring robust security measures against anomalies and cyber threats is foremost. This study advantages a thorough dataset derived from simulated WSN environments [11], [12] to mirror real network behaviors and possible security vulnerabilities.

2.2. Method

This part details the method utilized to discover and understand trends, and data based on dataset. An experimental methodology was used which involved three phases: exploratory data analysis (EDA), Modeling and design, and experiments. The methodology embraces an EDA to reveal implicit patterns and anomalies within the network data, data collection, preprocessing, feature engineering, model training, and evaluation. The first part focuses on understanding trends, and data through EDA, including time series, correlation, distribution, attack type analysis, and anomaly detection. The second part, based on the insights from EDA, involves the design and modeling of the ML and AI models for WSN. This understanding serves the feature engineering phase, where raw data attributes are transformed into features beneficial to machine learning algorithms. Experiments were run using Google Colab and Jupyter Notebook environments, benefiting their capabilities for data analysis, visualization, and ML model development. A thorough dataset obtained from simulated WSN environments was used, including several features (time, energy consumption, communication metrics, and environmental data).

2.3. Data preparation and feature engineering

Data preparation and feature engineering are important phases that ground the raw data with advanced machine learning algorithms. The process started with thorough data cleaning to remedy disparities and treat irrelevant entries. Consecutive cleaning and normalization techniques were utilized to scale numerical data, assuring no attribute unequal influenced the model. Missing values were systematically addressed by imputation, ensuring the dataset's integrity. Categorical variables were encoding, transforming

textual data into a numerical format compliant with algorithmic processing. This included transforming the 'Attack type' variable into an encoded format, crucial for classification tasks within cybersecurity domains.

Temporal feature extraction constituted an outstanding portion of the feature engineering process. Timestamp data were decomposed into discrete components such as hours, days, and months, providing granular insights into temporal patterns and anomalies. This temporal dissection allowed the models to notice patterns related to specific times, enhancing their predictive accuracy for time-sensitive anomalies. The dataset was split into training and testing datasets. The training dataset contains 80% (299,728 rows) and the testing dataset contains 20% (74,933 rows) of the total data. Data values were stored in the database and then exported in CSV files to apply the ML models to the data.

Furthermore, advanced feature engineering techniques were engaged to distill informative attributes from raw network metrics. Communication ratios, energy efficiency indicators, and network load metrics were calculated, transforming raw sensor data and network metrics into features with elevated predictive capabilities. For this study, key features integrated into model algorithms included:

- Temporal features: 'Hour' and 'DayOfWeek' extracted from the 'Time' data, offering insights into patterns that may correlate with network security incidents at specific times;
- Communication ratios: 'ADV_Ratio' and 'JOIN_Ratio', derived from the advanced-to-received and join-request-to-received message ratios, respectively, highlighting communication behavior anomalies;
- Distance metrics: 'Total_Dist_To_BS', combining 'Dist_To_CH' (distance to cluster head) and 'dist_CH_To_BS' (distance from cluster head to base station), to assess the impact of node positioning on network vulnerability;
- Energy efficiency indicators: 'Energy_per_Packet' and 'Energy_per_Dist', focusing on the energy consumed per data packet sent and per distance unit, which can signal inefficient or compromised nodes;
- Environmental changes: 'Temp_Change' and 'Humidity_Change', tracking abrupt variations in temperature and humidity that could affect sensor performance and potentially indicate tampering or anomalies;
- Network load: calculated as 'Network_Activity' and represented by 'Network_Load', providing a measure of the network's operational burden and identifying potential stress or attack vectors.

These engineered features outline the complex dynamics of WSNs, offering a nuanced understanding of network behavior that supports effective anomaly detection. This approach not only supports the instant detection of potential security breaches but also equips network administrators with actionable insights, thereby aiding proactive measures to reinforce network defenses.

2.4. Machine learning models

In this work, two machine learning algorithms were used for their adaptability and effectiveness in tackling classification challenges within the cybersecurity framework of WSNs. DT, and RF were selected to evaluate the best prediction model for WSN. These supervised learning ML algorithms are especially suited due to their capability to handle the dataset's complexity, which includes distinct features like temporal information, communication metrics, and environmental factors—all crucial for detecting anomalies. DT provides a model that refines the understanding of how various features impact the prediction of security threats. RF is a prediction method that extends this by aggregating multiple trees to enhance prediction accuracy and robustness against overfitting, ensuring the model's reliability across different WSN scenarios.

2.4.1. Algorithm 1: DT

At the center of our model selection is the DT algorithm, preferred for its simplicity and interpretability. This model constructs a tree-like structure of decisions, where each node represents a feature in the dataset, and each branch symbolizes a decision rule leading to different outcomes. In the context of WSNs, the DT provides an inherent mechanism to dissect and understand the underlying factors contributing to network anomalies. By thoroughly segmenting the dataset based on feature values, it gives insights into the significance of specific features in predicting security breaches. This approach not only supports anomaly detection but also facilitates the identification of potential areas for network security improvement.

2.4.2. Algorithm 2: RF

Building upon the DT foundation, the RF algorithm incorporates joint DT to form a packed model, significantly enhancing the predictive performance and stability. By generating a multitude of trees and aggregating their predictions, RF mitigates the overfitting issues commonly associated with single DT. It gains dominant accuracy through majority voting or averaging, making it particularly skilled at handling the complex and dynamic nature of cybersecurity threats within WSNs. This algorithm's capacity to analyze extensive datasets and notice complex patterns makes it an invaluable tool for preventive identifying and mitigating potential security incidents.

2.4.3. Feature importance and model demonstration

An essential component of utilizing these algorithms is the evaluation of feature importance, which uncovers the relative significance of each feature in the predictive models. This analysis not only improves the understanding of the models' decision-making processes but also tailors future data collection and preprocessing efforts. By identifying which features most forcibly influence the model's predictions, researchers can focus their efforts on the most applicable data, optimizing the efficiency and efficacy of cybersecurity measures in WSNs. Moreover, the upcoming phase was to train an ML model to dynamically classify and predict various types of anomalies and security threats within WSNs.

2.5. Models validation

Model validation was conducted through a series of performance evaluations using the standard model performance metrics such as accuracy, precision, recall, and F1 score. The behavior and performance of the model were assessed by using these metrics and their relevance to DT and RF. Accuracy measures the proportion of correctly classified instances out of the total instances evaluated [13], [14]. In the case of DT and RF models, accuracy reflects the overall correctness of the model's predictions, indicating how well they classify both normal and anomalous network activities. Precision quantifies the accuracy of positive predictions made by the model [15]. In our case, precision evaluates the DT and RF models' capacity to correctly identify actual network threats without falsely labeling normal activities as anomalies. Recall measures the model's ability to capture all positive instances in the dataset [16]. In the context of cybersecurity, recall assesses how fine the DT and RF models detect real network threats, ensuring minimal false negatives. The F1 score is the mean of precision and recall, providing a balanced assessment of a model's performance. It considers both false positives and false negatives, making it an applicable metric for assessing model effectiveness in imbalance scenarios [17]. The F1 score provides a thorough evaluation of DT and RF models, indicating robustness in detecting network anomalies while minimizing misclassifications. These metrics provide insights into the DT and RF model's performance, emphasizing their effectiveness in detecting anomalies within WSN.

3. RESULTS AND DISCUSSION

The analysis of the results highlights the pivotal role of feature engineering in optimizing model performance. After collecting the data, Google Colab and Jupyter Notebook were utilized for data analysis and visualization of data. In this section, we introduce the findings from our experiments and outline conclusive remarks based on the conducted analyses.

3.1. Exploratory data analysis (EDA)

3.1.1. Overview of dataset characteristics

Our analysis begins with an EDA, explaining key dataset characteristics, anomalies, and patterns. This inspection shows valuable insights into the distribution, variability, and relations among variables. The dataset utilized in this study demonstrates a wide variety of features that capture the intricate dynamics of wireless sensor networks (WSNs). By leveraging these diverse features, the analysis ensures a comprehensive understanding of the dataset, enabling the design of effective machine learning models for anomaly detection.

3.1.2. Time series analysis

Our primary aim of this phase was to reveal temporal patterns and potential anomalies embedded within critical variables (time, expanded energy, temperature, and humidity). We initiated the EDA with data cleaning and preparation steps, handling of missing values, and standardization of column names, to ensure dataset integrity. Following this, we engaged statistical summaries to highlight key statistical attributes such as central tendencies, dispersions, and ranges. The examination revealed no missing values in critical fields, thereby ensuring the completeness of our dataset. The statistical summaries bring significant variability within expanded energy (mean of ~0.306 units, standard deviation of 0.669 units), temperature (mean of ~15 °C, standard deviation of 14.444 °C), and humidity (mean of ~50%, standard deviation of 28.852%). Figure 1 shows the time series analysis data, visualizing expanded energy, temperature, and humidity over time. The presence of significant variability and no missing values in key fields indicates a solid and thorough dataset, vital for accurate anomaly detection.

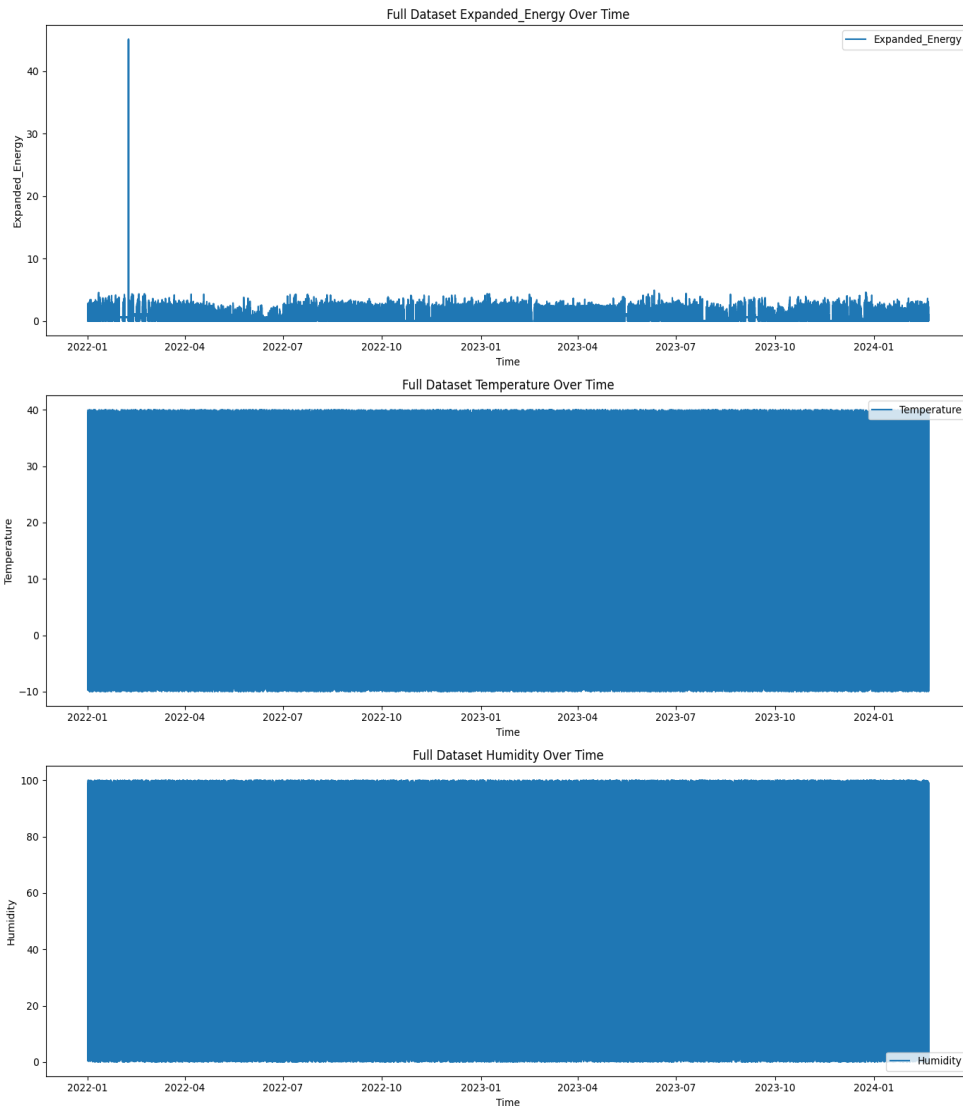


Figure 1. Current times series data

3.1.3. Correlation analysis

This phase pursued to explain the complex relationships between various features, crucial for our subsequent feature selection. A correlation matrix visually represents the Pearson correlation coefficients, which measure the linear relationship between the dataset's variables. Temperature and humidity emerged as insignificant correlations with other variables, indicating their potential independence or lack of influence within the network behavior captured by our dataset. However, expanded energy showed a moderate positive correlation with `dist_CH_To_BS` (0.38), implying escalated energy consumption with greater distances from the cluster head to the base station. Additionally, the correlation coefficients between `Dist_To_CH` and communication metrics such as `JOIN_S` (0.55) and `JOIN_R` (-0.16) suggest moderate correlations. This implies that nodes positioned farther from the cluster head exhibit varying communication behaviors, potentially influenced by signal range limitations or network congestion. Conversely, `dist_CH_To_BS` has a strong negative correlation with `SCH_R` (-0.68), potentially indicating scheduling communication issues with increased distance from the base station [18]. Figure 2 shows the correlation of the variables, highlighting key relationships between network metrics which are critical for understanding the dynamics of the WSN. These correlations hand vital insights into how network topology and communication behaviors collision energy consumption and scheduling, lead feature selection for successful anomaly detection.

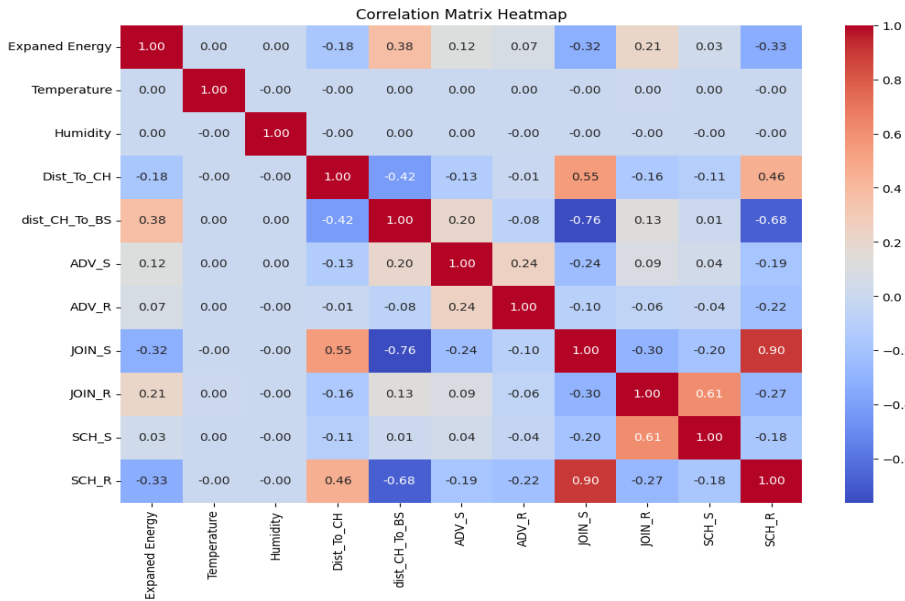


Figure 2. Pearson correlation between pairs of variables

3.1.4. Distribution analysis

This segment strived to examine the distributions of expanded energy, temperature, and humidity. Figure 3 shows the frequency distributions of data points within specified intervals. Expanded energy's distribution emerged skewed, intensified by a significant peak at lower values, indicating low-energy occurrences punctuated by infrequent high-energy anomalies or normal operation. Difference, temperature, and humidity histograms exhibited distinct values, implying potential data categorization. The skewed distribution of expanded energy put forward a greater number of normal operations with periodic anomalies, which is vital for training models to recognize infrequent but significant events.

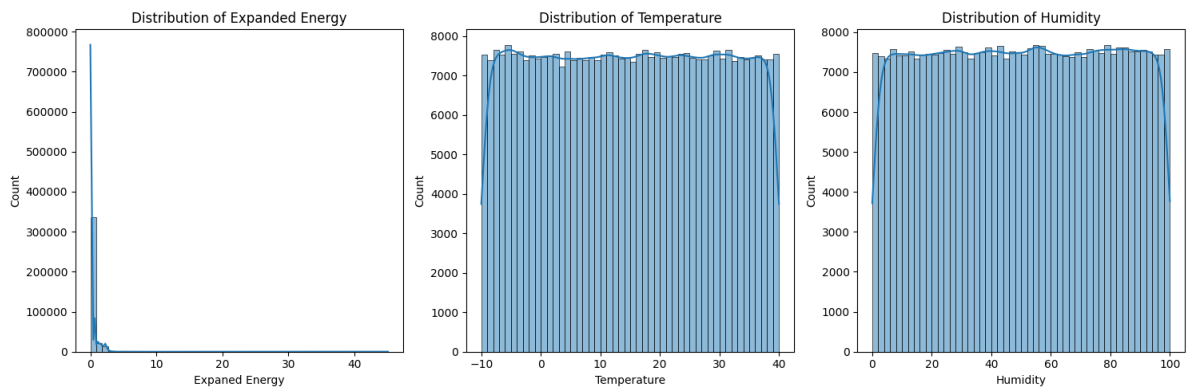


Figure 3. Plot of the frequency of data points that fall within specified ranges

3.1.5. Attack type analysis

This analysis aimed to explore the distribution of attack types within our dataset. Figure 4 shows a bar chart offering insights into the frequency distribution of distinct attack types. The distribution exhibits a stark difference, with 'Normal' attacks significantly outweighing other attack types, indicating that most of the data points are labeled as normal. The other attack types such as 'Grayhole', 'Blackhole', 'TDMA', and 'Flooding' occur much less frequently [19]. The imbalance in attack types emphasizes the need for models that can handle imbalanced datasets, concentrating on the infrequent but critical instances of network attacks.

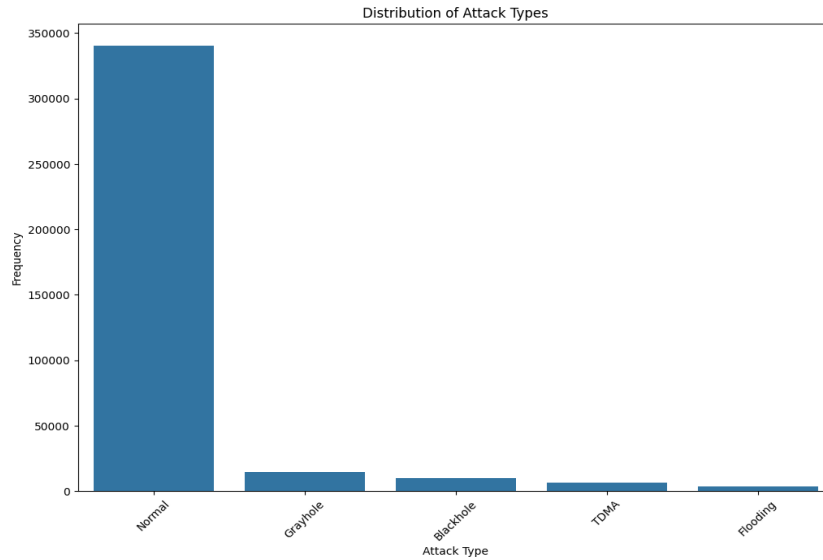


Figure 4. Visualizing the frequency distribution of attack-type data

3.2. Performance of machine learning models

3.2.1. DT model performance

The evaluation of the DT model is directed to assess its effectiveness in classifying WSN attack types, aligning with the research objective of accurate anomaly detection within WSNs. Using features engineered from the dataset, including temporal features obtaining cyclical patterns and communication ratios reflecting network efficiency, the DT model was trained to notice patterns indicative of security incidents. The evaluation revealed performance metrics with an accuracy of 99.47%, precision of 96.53%, recall of 97.26%, and F1 score of 96.89%. These values reflect the model's ability to accurately classify instances, minimize false positives, and capture genuine security incidents effectively. Figure 5 shows the decision boundaries of the DT model, by mapping between attack types and encoded labels within the dataset. Figure 6 presents the feature importance plot for the DT model and uncovers the significance of network communication metrics (ADV_S, 0.5) and energy consumption (expanded energy, 0.2-0.3) in the classification process. Lower-importance features, such as SCH_S (<0.1), still contribute to the model's decision-making, emphasizing the critical role of network communication metrics and energy consumption in identifying various network attack types. These insights give valuable advice for network administrators and security professionals in effectively monitoring and mitigating security threats within WSN [20].

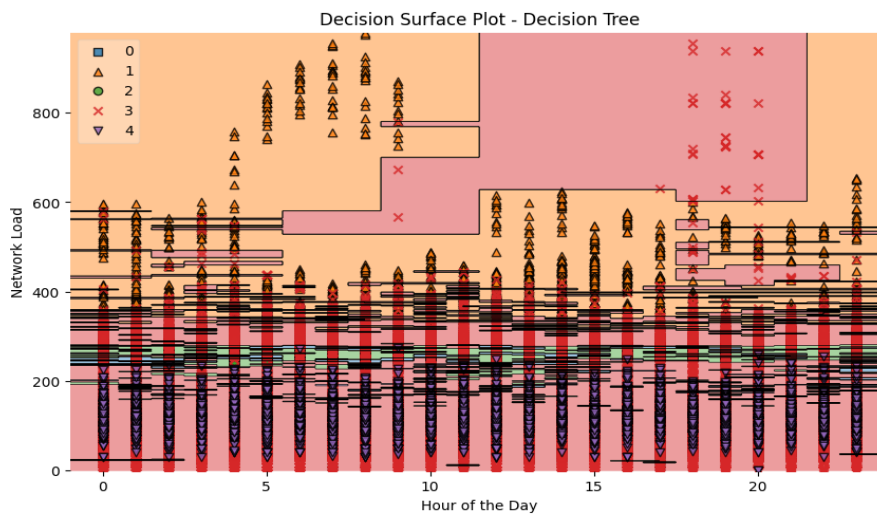


Figure 5. DT decision surface plot of attack types: blackhole 0, flooding 1, grayhole 2, normal 3, and TDMA 4

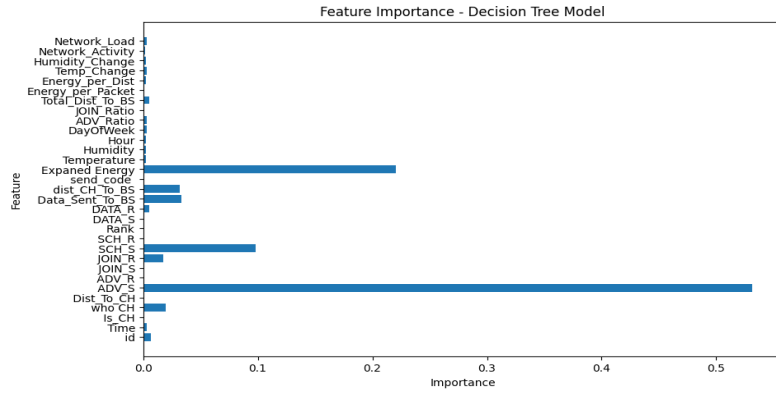


Figure 6. Feature importance-DT model

3.2.2. RF model performance

The evaluation of the RF model focused on assessing its classification performance in WSN anomaly detection, using the engineered dataset features. By aggregating predictions from multiple DT, the RF model used a variety of individual trees to enhance predictive accuracy and generalization. Features such as distance metrics and environmental anomalies were essential in training the model to detect subtle patterns indicative of security incidents [21]. The RF model showed dominant performance metrics compared to the DT model, with an accuracy of 99.71%, precision of 98.32%, recall of 98.22%, and F1 score of 98.23%. These values emphasize the model's advanced ability to precisely classify instances and minimize false alarms, further aiding its efficacy in WSN anomaly detection. Table 1 shows the feature importance of the RF model, indicating that expanded energy (0.512) and ADV_S (0.488) have a relatively stronger influence on the model's predictions [22].

Table 1. Feature importance table-RF model

Feature ID	Feature	Value
0	Expanded energy	0.51209
1	ADV_S	0.4879

Figure 7 shows the most important features for classifying network security incidents in the RF model, with expanded energy (11.72%), ADV_S (11.05%), and Total_Dist_To_BS (8.18%) contributing significantly to the model's decision-making process. Metrics, (Data_Sent_To_BS, and SCH_S), play vital roles in analyzing network traffic, behavior, and efficiency.

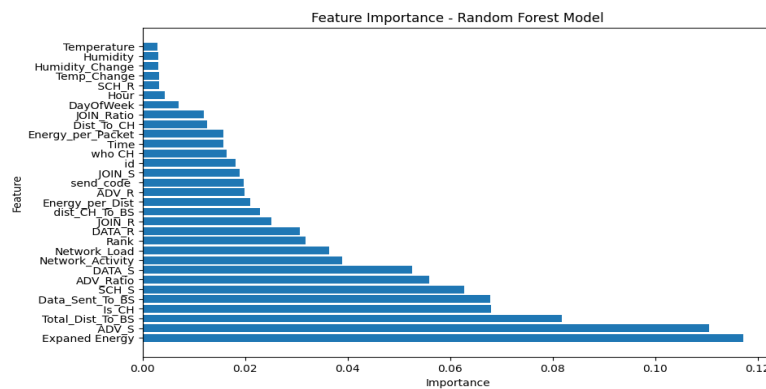


Figure 7. Feature importance-RF model

3.3. Comparative analysis

3.3.1. Performance comparison between decision tree and RF

To confirm the prediction ability of the proposed models, the accuracy of the DT and RF models was evaluated based on their performance (accuracy scores). The inputs to both models included network security incidents, network communication metrics, environmental variables, and other useful parameters

extracted from sensor readings or network logs [23]. Figure 8 shows the prediction results from comparing the DT and RF models. DT model realized an accuracy of 99.47%, while the RF model outperformed it with an accuracy of 99.71%. This means that both models performed exceptionally well in classifying network security incidents in our dataset. Still, the RF model demonstrated a slightly higher accuracy, indicating that its ensemble learning approach, which combines multiple DT, may have contributed to improved classification accuracy compared to the single DT approach of the DT model [24].

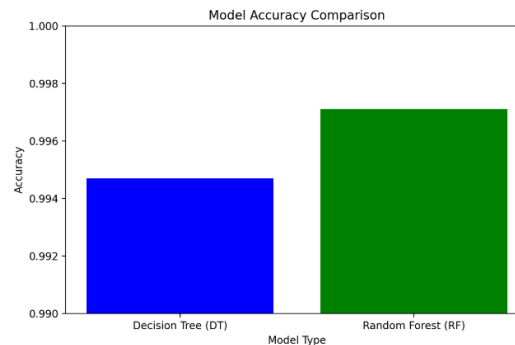


Figure 8. Comparison of the performance metrics

The results of this study reinforce the effectiveness of using ML techniques, DT, and RF models, in detecting network security incidents in WSNs. These findings indicate that network administrators can implement our framework to identify and mitigate security threats in real-time, as a result enhancing the security posture of WSN infrastructures. While the RF model has proven just higher accuracy than the DT model, it was stunning to observe that some features (energy consumption and network load), played a more key role in the classification process than originally expected, indicating that future anomaly detection systems should consider these factors to advance detection rates. This study has limitations the actual models do not consider the possible impact of advancing cyber threats and flexible attack strategies, requiring continuing updates and re-training to preserve effectiveness. Future research should center on the design and implementation of a sample WSN secure communication infrastructure, integrating both hardware and software, to run experimental studies of the proposed solutions in practical implementation areas [25].

4. CONCLUSION

Our study unveils the effectiveness of machine learning algorithms, particularly RF, in detecting network security incidents in WSNs. By using engineered features and visualization, we see important insights into the basic patterns and decision logic of the models. The proposed model will assist network administrators and security analysts with an efficient mechanism for detecting and mitigating network security incidents in WSNs, enhancing the security posture of IoT deployments. This study confirmed that using machine learning techniques, such as DT and RF, can advance anomaly detection accuracy in WSNs, regarding timely threat identification and response. The future work of WSN security interests the implementation of a sample WSN secure communication infrastructure, letting experimental studies assess the effectiveness of proposed security models. By focusing on practical implementation, experimental evaluation, and industry cooperation, we can enhance the evolution of well and resilient security models for WSNs, contributing to strengthening the cybersecurity of IoT deployments.

ACKNOWLEDGEMENTS

This research has been funded by the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan (Grant No. AP19680345). We thank the institutions for the support of funding.



REFERENCES

- [1] S. Damadam, M. Zourbakhsh, R. Javidan, and A. Faroughi, "An intelligent iot based traffic light management system: deep reinforcement learning," *Smart Cities*, vol. 5, no. 4, pp. 1293–1311, Sep. 2022, doi: 10.3390/smartcities5040066.
- [2] H. Benaddi, K. Ibrahim, A. Benslimane, and J. Qadir, "A deep reinforcement learning based intrusion detection system (drl-ids) for securing wireless sensor networks and internet of things," in *Lecture Notes of the Institute for Computer Sciences, Social- Informatics and Telecommunications Engineering, LNICST*, vol. 317 LNICST, 2020, pp. 73–87. doi: 10.1007/978-3-030-52988-8_7.




- [3] J. B. Awotunde and S. Misra, "Feature extraction and artificial intelligence-based intrusion detection model for a secure internet of things networks," in *Lecture Notes on Data Engineering and Communications Technologies*, vol. 109, 2022, pp. 21–44. doi: 10.1007/978-3-030-93453-8_2.
- [4] A. Kavousi-Fard, W. Su, and T. Jin, "A machine-learning-based cyber attack detection model for wireless sensor networks in microgrids," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 1, pp. 650–658, Jan. 2021, doi: 10.1109/TII.2020.2964704.
- [5] L. Njilla, L. Pearlstein, X. W. Wu, A. Lutz, and S. Ezekiel, "Internet of things anomaly detection using machine learning," in *Proceedings - Applied Imagery Pattern Recognition Workshop*, IEEE, Oct. 2019, pp. 1–6. doi: 10.1109/AIPR47015.2019.9174569.
- [6] A. Diro, N. Chilamkurti, V. D. Nguyen, and W. Heyne, "A comprehensive study of anomaly detection schemes in iot networks using machine learning algorithms," *Sensors*, vol. 21, no. 24, p. 8320, Dec. 2021, doi: 10.3390/s21248320.
- [7] I. S. Thaseen, V. Mohanraj, S. Ramachandran, K. Sanapala, and S. S. Yeo, "A hadoop based framework integrating machine learning classifiers for anomaly detection in the internet of things," *Electronics (Switzerland)*, vol. 10, no. 16, p. 1955, Aug. 2021, doi: 10.3390/electronics10161955.
- [8] J. Jiang, G. Han, L. Liu, L. Shu, and M. Guizani, "Outlier detection approaches based on machine learning in the internet-of-things," *IEEE Wireless Communications*, vol. 27, no. 3, pp. 53–59, Jun. 2020, doi: 10.1109/MWC.001.1900410.
- [9] I. Almomani, B. Al-Kasasbeh, and M. Al-Akhras, "WSN-DS: a dataset for intrusion detection systems in wireless sensor networks," *Journal of Sensors*, vol. 2016, pp. 1–16, 2016, doi: 10.1155/2016/4731953.
- [10] M. Al Samara, I. Bennis, A. Abouaissa, and P. Lorenz, "A survey of outlier detection techniques in iot: review and classification," *Journal of Sensor and Actuator Networks*, vol. 11, no. 1, p. 4, Jan. 2022, doi: 10.3390/jsan11010004.
- [11] M. S. Alsahli, M. M. Almasri, M. Al-Akhras, A. I. Al-Issa, and M. Alawairdhi, "Evaluation of machine learning algorithms for intrusion detection system in WSN," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 5, pp. 617–626, 2021, doi: 10.14569/IJACSA.2021.0120574.
- [12] S. E. Quincozes and J. F. Kazienko, "Machine learning methods assessment for denial of service detection in wireless sensor networks," in *IEEE World Forum on Internet of Things, WF-IoT 2020 - Symposium Proceedings*, IEEE, Jun. 2020, pp. 1–6. doi: 10.1109/WF-IoT48130.2020.9221146.
- [13] R. Alshinina and K. Elleithy, "A highly accurate machine learning approach for developing wireless sensor network middleware," in *Wireless Telecommunications Symposium*, IEEE, Apr. 2018, pp. 1–7. doi: 10.1109/WTS.2018.8363955.
- [14] I. G. A. Poornima and B. Paramasivan, "Anomaly detection in wireless sensor network using machine learning algorithm," *Computer Communications*, vol. 151, pp. 331–337, Feb. 2020, doi: 10.1016/j.comcom.2020.01.005.
- [15] "COMCOM 2017 organizing chairs," in *2017 International Conference on Applied Computer and Communication Technologies (ComCom)*, IEEE, May 2017, pp. 1–3. doi: 10.1109/comcom.2017.8167078.
- [16] P. Radivojac, U. Korad, K. M. Sivalingam, and Z. Obradovic, "Learning from class-imbalanced data in wireless sensor networks," in *IEEE Vehicular Technology Conference*, IEEE, 2003, pp. 3030–3034. doi: 10.1109/vetecf.2003.1286180.
- [17] M. Owusu-Adjei, J. Ben Hayfron-Acquah, T. Frimpong, and G. Abdul-Salaam, "Imbalanced class distribution and performance evaluation metrics: A systematic review of prediction accuracy for determining model performance in healthcare systems," *PLOS Digital Health*, vol. 2, no. 11 November, p. e0000290, Nov. 2023, doi: 10.1371/journal.pdig.0000290.
- [18] A. Srivastava and M. R. Bharti, "Hybrid machine learning model for anomaly detection in unlabelled data of wireless sensor networks," *Wireless Personal Communications*, vol. 129, no. 4, pp. 2693–2710, Apr. 2023, doi: 10.1007/s11277-023-10253-2.
- [19] R. Ahmad, R. Wazirali, and T. Abu-Ain, "Machine learning for wireless sensor networks security: an overview of challenges and issues," *Sensors*, vol. 22, no. 13, p. 4730, Jun. 2022, doi: 10.3390/s22134730.
- [20] O. S. Egwuche, A. Singh, A. E. Ezugwu, J. Greeff, M. O. Olusanya, and L. Abualigah, "Machine learning for coverage optimization in wireless sensor networks: a comprehensive review," *Annals of Operations Research*, Nov. 2023, doi: 10.1007/s10479-023-05657-z.
- [21] M. A. Alsheikh, S. Lin, D. Niyato, and H. P. Tan, "Machine learning in wireless sensor networks: Algorithms, strategies, and applications," *IEEE Communications Surveys and Tutorials*, vol. 16, no. 4, pp. 1996–2018, 2014, doi: 10.1109/COMST.2014.2320099.
- [22] E. Ancillotti, C. Vallati, R. Bruno, and E. Mingozzi, "A reinforcement learning-based link quality estimation strategy for RPL and its impact on topology management," *Computer Communications*, vol. 112, pp. 1–13, Nov. 2017, doi: 10.1016/j.comcom.2017.08.005.
- [23] G. M. Gandhi, S. Parthiban, N. Thummalu, and A. Christy, "Ndvi: vegetation change detection using remote sensing and gis - a case study of vellore district," *Procedia Computer Science*, vol. 57, pp. 1199–1210, 2015, doi: 10.1016/j.procs.2015.07.415.
- [24] V. Sivagaminathan, M. Sharma, and S. K. Henge, "Intrusion detection systems for wireless sensor networks using computational intelligence techniques," *Cybersecurity*, vol. 6, no. 1, p. 27, Oct. 2023, doi: 10.1186/s42400-023-00161-0.
- [25] H. M. Ammari, "Coverage in wireless sensor networks: a survey," *Network Protocols and Algorithms*, vol. 2, no. 2, Jun. 2010, doi: 10.5296/npa.v2i2.276.

BIOGRAPHIES OF AUTHORS






Prof. Dr. Tamara Zhukabayeva    received the Ph.D. degree from Satbayev University, Kazakhstan. She is currently an Associate Professor in informatics, computer engineering and management with L. N. Gumilyov Eurasian National University, Astana, Kazakhstan. She is also an Associate Member of the Universal Association of Computer and Electronics Engineers, has membership in scientific societies in The Society of Digital Information and Wireless Communications (SDIWC) and Universal Association of Computer and Electronics Engineers. She has published over 70 scientific and educational-methodical works: in the Republic of Kazakhstan, and in countries of far and near abroad, including a foreign edition from the Clarivate Analytics Database, Scopus. She is the author and coauthor of educational publications and scientific monographs, has an innovative patent and copyright certificates for intellectual property rights. She can be contacted at email: tamara_kokenovna@mail.ru.






Atdhe Buja PhD (c)    is an Associate Professor at Commonwealth University of Pennsylvania Department of Math, Digital Forensics, USA. He holds an MSc degree in Computer Science, and certified professional in the industry as a Certified EC-Council Instructor (CEI), CEH, MCITP, OCA, and CIO. His research areas are cybersecurity for IoT, IIoT, WSN, and digital forensics. He is the director of the ICT Academy Research and Innovation Lab, and member of the International Science Complex in Astana, and a NIST GCTC member. He is the founder of ICT Academy which is a Cyber Security-based company and their innovative services received appreciation at national and international levels. He can be contacted at email: atdhe.buja@academyict.net.



Dr. Melinda Pacolli    PhD in IT Management, specialised in ERP Systems and digital transformation. Currently engaged as Associate Professor in Informarion Systems, Data Science and Business Analytics Programme. Also active in the IT industry, with over 15 years of professional experience in implementation of ERP systems and in providing digital transformation consultancy to SMEs in different countries, Italy, Switzerland, Macedonia and latest in Kosovo. She can be contacted at email: pacollimelinda@gmail.com.



Yerik Mardenov    Graduate of OP 6D070400 Computer technologies and software, Eurasian National University named after L. N. Gumilyov. The topic of the dissertation is “Development and research of algorithms and models for analyzing the security of software and hardware components of wireless sensor networks.” Director of the Information Technology Department at Astana International University. He can be contacted at email: emardenov@gmail.com.