# Tackling the anomaly detection challenge in large-scale wireless sensor networks

**Tamara Zhukabayeva[1,2,3], Aigul Adamova[1,3], Lyazzat Zholshiyeva[1], Yerik Mardenov[1,4], Nurdaulet Karabayev[1,2], Dilaram Baumuratova[1,4]**

[1]Institute of Information Technology and Security, International Science Complex ASTANA, Astana, Kazakhstan
[2]Faculty of Information Technology, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan
[3]Department of Computer Engineering, Astana IT University, Astana, Kazakhstan
[4]School of Information Technology and Engineering, Astana International University, Astana, Kazakhstan

**Article Info**

***Article history:***

Received Jul 5, 2024
Revised Nov 4, 2024
Accepted Nov 20, 2024

***Keywords:***

Anomaly detection
K-nearest neighbor
Machine learning
Wireless sensor networks
Z score

**ABSTRACT**

One of the areas of ensuring the security of a wireless sensor network (WSN) is anomaly detection, which identifies deviations from normal behavior. In our paper, we investigate the optimal anomaly detection algorithms in a WSN. We highlight the problems in anomaly detection, and we also propose a new methodology using machine learning. The effectiveness of the k-nearest neighbors (KNN) and Z score methods are evaluated on the data obtained from WSN devices in real time. According to the experimental study, the Z score methodology showed a 98.9% level of accuracy, which was much superior to the KNN 43.7% method. In order to ensure accurate anomaly detection, it is crucial to have access to high-quality data when conducting a study. Our research enhances the field of WSN security by offering a novel approach for detecting anomalies. We compare the performance of two methods and provide evidence of the superior effectiveness of the Z score method. Our future research will focus on exploring and comparing several approaches to identify the most effective anomaly detection method, with the ultimate goal of enhancing the security of WSN.

***Corresponding Author:***

Aigul Adamova
Department of Computer Engineering, Astana IT University
55/11 Mangilik El, Astana IT University, 010000 Astana, Kazakhstan
Email: aigul.adamova@astanait.edu.kz

## 1. INTRODUCTION

Wireless sensor networks (WSN) represent modern solutions for interactions with the environment [1], [2]. Currently, as a result of the widespread use of WSN in a wide range of applications covering everyday life and industrial settings, the issues of ensuring the security of the information space are becoming increasingly complex. The report “global wireless sensor networks market” by research and markets highlights the impact of WSN on the development of various industries that rely on automation and data-driven decision making [3]. A WSN uses sensors to monitor and control various processes, predict equipment failures and optimize resource utilization [4], [5]. Ensuring the security of WSN is an urgent task. There are many vulnerabilities in WSN, at the same time they are vulnerable to various types of attacks, such as denial-of-service attacks, physical attacks, node replication attacks, and traffic analysis attacks. Since data transmission is carried out via wireless technologies, the attack can be carried out from various remote locations at any time [6], [7]. The task of developing and implementing reliable security measures, such as anomaly detection, plays an important role in maintaining trust in WSN [8]–[10]. Anomaly detection is a serious problem and does not lose its relevance to this day. They belong to many application areas, each of which has its own peculiarities and limitations, such as cybersecurity, failure prevention, industrial automation, preventive maintenance and much more [11]–[13]. The essence of the anomaly detection method is to study the normal events in the network. Anomaly detection detects anomalies on the basis of a predefined set of normal data. Hence, this type of outlier detection can detect even unknown attacks. Importantly, although anomaly detection is significantly high, it can also have a high false positive rate [14]–[16]. Anomaly detection in WSN faces challenges due to the dynamic nature of networks, resource constraints and the need to adapt to nonstationary data distributions.

The aim of this study is to investigate and determine the optimal algorithm for anomaly detection in WSN. The applied machine learning methods demonstrate a high percentage of anomaly detection in real-time sensor data. The present study makes several contributions: we developed a method for anomaly detection and presented in detail the various problems associated with using certain methods, we suggested a methodology for detecting anomalies, and we demonstrated the applicability of the k-nearest neighbors (KNN) and Z score methods in detecting anomalies. The remainder of the paper is organized as follows: section 2 presents the research methodology, section 3 discusses the anomalies and existing anomaly detection methods, and the proposed methodology and experimental study are presented in sections 4. Section 5 concludes the presented study.

## 2. RESEARCH METHODOLOGY

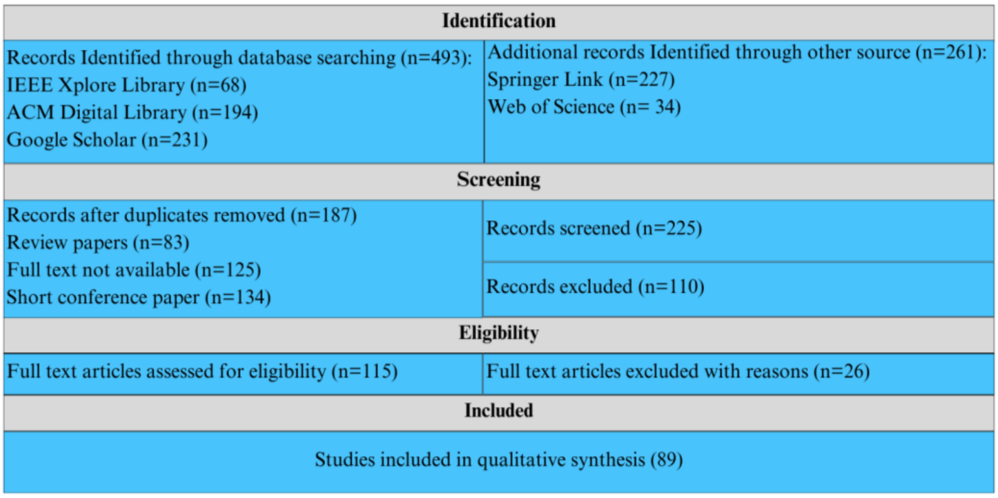
One of the initial objectives of the study was to provide up-to-date information on the direction of anomaly detection in WSN. This section provides information on the current state of anomaly detection research and highlights open issues. In this context, we conducted a comprehensive review of works from 2020 to 2024. During the review of scientific papers, the applied methods, techniques, and algorithms for anomaly detection in WSN were considered in relation to the following questions:

Q1: Identify the main characteristics of anomaly detection methods in WSN?

Q2: What metrics are used to evaluate anomaly detection methods in WSN?

Q3: In which areas are the proposed algorithms for WSN anomaly detection applicable?

The systematic review was based on papers that were obtained from databases such as the IEEE Xplore Library, ACM Digital Library, Scopus, Springer Link, Elsevier, and Web of Science. The papers for the period from 2020–2024 were analyzed, and papers that were review papers, papers that did not have access to the full text and short conference papers were excluded. The search was conducted via the keywords “anomaly detection,” “wireless sensor network,” “machine learning,” “outlier detection,” and “intrusion detection.” The stages of the systematic review are depicted in Figure 1, which was prepared on the basis of PRISMA. As a result, 89 research papers were selected for the systematic review.



Identification

Records Identified through database searching (n=493):
IEEE Xplore Library (n=68)
ACM Digital Library (n=194)
Google Scholar (n=231)

Additional records Identified through other source (n=261):
Springer Link (n=227)
Web of Science (n= 34)

Screening

Records after duplicates removed (n=187)
Review papers (n=83)
Full text not available (n=125)
Short conference paper (n=134)

Records screened (n=225)

Records excluded (n=110)

Eligibility

Full text articles assessed for eligibility (n=115)

Full text articles excluded with reasons (n=26)

Included

Studies included in qualitative synthesis (89)

Figure 1. PRISMA flowchart of paper selection

Anomaly detection is an important research area and is the focus of many research papers. Machine learning techniques are actively applied in anomaly detection, as evidenced by the systematic review conducted [17]. Bacha *et al.* [18] proposed an intrusion detection system to prevent a wide range of cyber-attacks in internet of things (IoT) environments. The proposed system uses the kernel principal component analysis method to minimize the dimensionality of data features and improve the efficiency of anomaly detection. An anomaly is detected via traffic analysis via an extreme kernel learning machine [18].

Bhatia and Sangwan [19] have identified an approach that helps to mitigate real-time abuse based on the IoT in a proactive, reactive or predictive manner. Inuwa and Das [20] demonstrated the use of various machine learning techniques to detect cyber anomalies in IoT systems. Dissem *et al.* [21] proposed a reconstruction-based anomaly detection system using autoencoders in which we trained the model on anomaly free samples by minimizing the error between the original and reconstructed sequences. Affane *et al.* [22] highlighted the importance of anomaly detection and proposed a new method to analyze and classify WSN datasets. The method proposed by the authors is based on stochastic models incorporating predictive assumptions. To demonstrate the effectiveness of the proposed approach, the authors compare it with the support vector machines (SVM), naive Bayes (NB), decision tree (DT) and random forest (RF) methods.

Algarni *et al.* [23] investigated wireless communication technologies for maritime activities and highlighted the maintenance of safety through anomaly detection. The authors propose a solution based on edge computing and machine learning techniques such as long short-term memory and isolation forest (IF). Alangari [24] proposed a hybrid optimization method using unsupervised machine learning techniques. The sensors of WSN are protected from anomalies and various attacks by forming groups on the basis of secure certificates and trust filtering via the K-means method. Sivagaminathan *et al.* [25] presented an approach to detect malicious network connections on the basis of data mining and machine learning methods KNN, artificial neural network (ANN), and DT which were used with particle swarm optimization (PSO) selection features. As a result, the ANN and PSO classification methods have achieved good results. Moundounga *et al.* [26] have demonstrated a new method to analyze and classify a WSN dataset to improve its security via machine learning techniques. The proposed models are based on stochastic assumptions of the hidden Markov model and Gaussian mixture model. Srivastava and Bharti [27], in their publication, proposed a hybrid model of a single class SVM and IF. Anomaly detection is performed in two steps: the first step is conversion to labeled data, and the second step is the anomaly detection process. The results of the study are impressive.

By analyzing the selected papers one can easily notice that machine learning methods such as KNN and Z score analysis are not so common [28]. Therefore, the experimental study of these methods is useful and will help achieve this objective. Despite the amount of work done and achievements made by researchers, there are still open questions concerning how to ensure high accuracy in anomaly detection. The process of minimizing false positives is an important issue in WSN with a variety of data types. Low accuracy in anomaly detection can lead to serious risks. Therefore, a study to improve the reliability of anomaly detection algorithms via KNN and Z score machine learning techniques can help improve accuracy.

## 3. METHOD

WSN data are critical for good decision-making, so detecting anomalies that may indicate sensor malfunctions, security threats or unexpected events is important. There are different methods for detecting anomalies in WSN. While the generic approach works by examining what “normal” data look like and then flagging anything that deviates significantly, the supervised approach requires labeled data. This section discusses the different types of anomalies and presents an analysis of the methods used in practice to detect anomalies in WSN.

### 3.1. Anomalies

WSN consists of multiple nodes distributed over a certain area. WSN have the ability to ingest data, perform computations, coordinate complete activities and simultaneously transmit data to users. Depending on the deployment area, various design constraints, such as communication barriers, computational power and energy consumption, exist [29]. The quality of the data received from WSN nodes affects decision-making; hence, studying different types of anomalies is an important task [30]. Figure 2 summarizes the types of WSN anomalies that can adversely affect the important decision-making process, and examples of anomalies are also given.

### 3.2. Methods for anomaly detection in WSN

The anomaly detection method is reduced to the study of normal events, the deviation of which is used to detect anomalies. Since this method detects any changes, it may well detect unknown attacks, but at the same time, there is a probability of false positives [31]. Anomaly detection allows the identification of unauthorized actions in the form of various attacks. As an attack is defined as a certain sequence of actions, by applying them to the fields of the identified object, information regarding its affiliation with this attack can be obtained. Figure 3 shows the network anomaly detection scheme, where network traffic is used as the original data. The network packets' collected properties are forwarded to a module that examines and verifies the input data against predefined rules. If any of these rules are activated, it raises an alarm for a potential threat. A key problem in the design of any anomaly detection system is the effective design of the rule assignment mechanism [32].

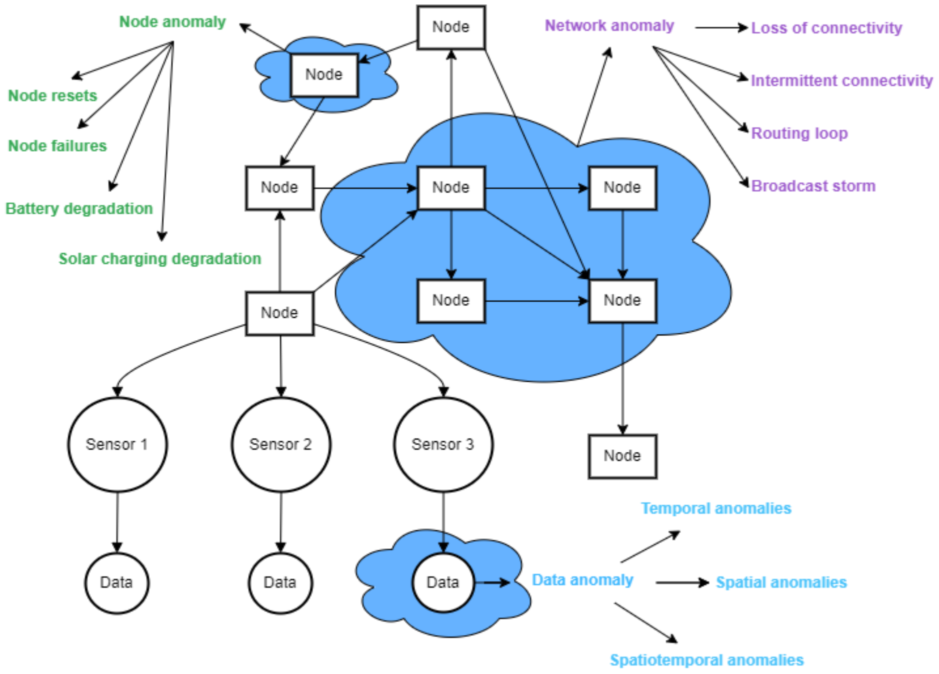


Figure 2. Types of WSN anomaly

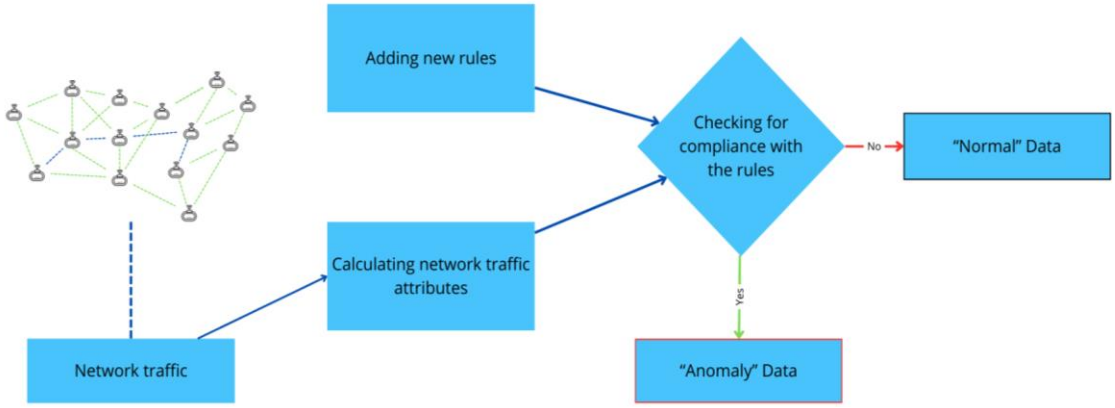

Adding new rules
Checking for compliance with the rules
No
"Normal" Data
Calculating network traffic attributes
Yes
Network traffic
"Anomaly" Data

Figure 3. Network anomaly detection scheme

There are three approaches to anomaly detection: supervised, unsupervised and semi supervised anomaly detection, Figure 4. Supervised anomaly detection methods use a labeled dataset that trains a classifier. The results of the studies by Castellani *et al.* [33] demonstrated the successful application of the supervised anomaly detection method in industrial settings. As noted in the work of Nassif *et al.* [34] presents 29 different machine learning models that were applied in anomaly detection experiments. As a result, the authors noted that unsupervised anomaly detection is used by researchers more often than other classification anomaly detection systems. Unsupervised anomaly detection methods detect anomalies in a test dataset that does not have any labels on the basis solely of the intrinsic properties of the data. It relies on statistical or distance-based measures to assess the difference from the rest of the data. Khan and Haroon [35] noted the features of unsupervised learning in the process of anomaly detection. Semi supervised anomaly detection methods use a regular labeled training dataset to build a model that represents normal behavior. They then use the built model to detect anomalies. Lu *et al.* [36] demonstrated the applicability of semi supervised learning to anomaly detection in cellular networks.

Existing anomaly detection methods include approaches such as classification, clustering, and static and artificial intelligence [37]–[40]. Research in this area has made significant contributions to anomaly detection via machine learning techniques such as KNN, RF, SVM, spatial clustering of applications with density-based noise (DBSCAN) and DT [41]–[44]. Table 1 presents a systematic description of the results of the anomaly detection work, indicating the method used, the dataset, and the results in the form of various metrics.

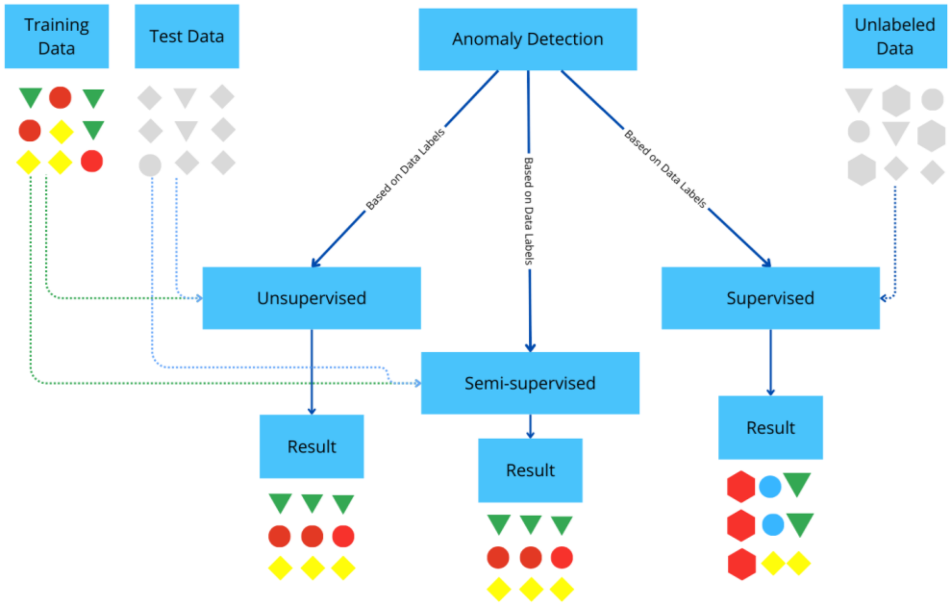


Figure 4. General types of anomaly detection techniques

Table 1. Systematic description of the results of the anomaly detection work

| Work | Method | Dataset | Metrics/performance |
|---|---|---|---|
| Poornima and Paramasivan [45] | An online identification based on linear weighted projective regression | Intel Berkeley Research Lab | Accuracy 0.91<br>Recall 0.86<br>Precision 0.85<br>F1-score 0.86<br>AUC 0.54 |
| Chen and Li [46] | An anomaly detection approach based on spatio-temporal and attribute correlations. | Grand-St-Bernard | N-STASVDD<br>AUC 0.99<br>RN-STASVDD<br>AUC 0.99 |
| Shi and Shen [47] | An approach based on artificial immune network | ISCX 2012 IDS<br>NSL-KDD | Accuracy 80.41%; 93.54% |
| SaiSindhuTheja and Shyam [48] | The oppositional crow search algorithm | KDD cup 99 | Accuracy 94.12%<br>Recall 95.13%<br>Precision 98,18%<br>F1-score 93.56% |
| Al-Turaiki and Altwaijry [49] | The convolutional neural network - deep feature synthesis | NSL-KDD | Accuracy 99.62%<br>Recall 99.6%<br>Precision 99.7%<br>F1-score 99.6% |
| Yao *et al.* [50] | A method based on principal component analysis and a deep convolutional neural network | KDDcup99<br>NSL-KDD<br>UNSW-NB15 | Accuracy 97.83%; 92.28%; 96.76%<br>Recall 97.83%; 92.28%; 96.76%<br>Precision 98.02%; 92.99%; 97.17%<br>F1-score 97.92%; 92.63%; 96.96% |
| Jain *et al.* [51] | Support vector machine | Testbed<br>NSL-KDD<br>CIDDS-2017 | Accuracy: 91.84%; 99.1%; 88.3%.<br>Recall: 94.30%; 99.2%; 91.7%.<br>F1-score: 92.9%; 99.15%; 89.6%. |

Summarizing the review and analysis work, we note that most of the methods demonstrate high accuracy and performance, particularly approaches that use deep neural networks and machine learning methods. Methods using complex models such as convolutional neural networks (CNN) and recurrent neural networks (RNN) have superior precision and few false alarms [49]. The reviewed studies use a variety of datasets, both synthetic and data read from real devices, which allows the evaluation of methods in different environments but also indicates the need for standardization for a fairer comparison [45]–[51]. Note that methods that consider spatiotemporal correlations and use hybrid approaches yield better results than traditional methods do [52], [53]. To date, different methods have been used in different studies, and Table 2 summarizes the advantages and disadvantages of the studied anomaly detection approaches.

Analyzing the advantages and disadvantages of the methods, we note that the choice of the method depends on the specifics of the problem, available resources and type of data [54]–[57]. Methods that require high computational resources (clustering, machine learning methods) provide high accuracy but may be unsuitable for sensors with limited resources [58]–[63]. Together, we note that machine learning and classification methods require careful tuning and training before being implemented in real-world applications, which increases the time and resources required for their application [60], [61]. In general, the selection of a suitable method for anomaly detection in WSN should be based on the balance between accuracy, computational cost, and adaptability to changes in the data.

Table 2. The advantages and disadvantages of the studied anomaly detection approaches

| Techniques | Advantages | Disadvantages |
|---|---|---|
| Statistical techniques [54], [55] | Identifies sensor malfunctions and anomalies in IoT using a probability distribution model. | Typically, IoT devices are used in real-world scenarios where there is no information about the distribution of sensor data and, accordingly, a parametric statistical approach is not useful. |
| | Sensor malfunctions and anomalies are identified using temporal correlation. Any unplanned changes in the distribution of data lead to a decline in temporal correlations, resulting in the detection of anomalies. | Non-parametric statistical models are often not suitable for intensive work with data obtained from IoT devices. |
| | | High computational costs for managing multidimensional data cannot be ruled out. |
| Nearest-neighbor techniques [56], [57] | Very easy to use with respect to various types of data that have been collected from IoT system sensors. | Dramatic increase in computational complexity when using multidimensional data. |
| | The data does not require any preliminary preparation. It is only necessary to correctly select the appropriate distance metric that will be used for their analysis. | Weak scalability, especially in the context of the IoT devices. |
| | | The sensor failures and anomaly detection have a significant issue with a high probability of false negatives. |
| Machine learning techniques [58], [59] | High applicability, the model generalizes the received data points and skips fragmented and noisy data. | The model needs to be refined and tested before being used in real-life scenarios. |
| | When new data arrives, retraining the model is not necessary | Difficulties in setting up and adapting the model when implementing it |
| Cluster techniques [60], [61] | Easy adaptability to incremental mode of operation, which requires only checking new data points for failures and anomalies, without requiring additional monitoring | High computational cost, inefficient in processing multidimensional data from sensors to detect faults |
| | High adaptability and easy integration of new data, which makes it applicable for detecting anomalies in data obtained from IoT devices. | High computational complexity, not practical for use with resource-constrained devices |
| Classification techniques [62], [63] | Applicability, the model can be used for multivariate data, but may be limited for other types of data. | High computational complexity |
| | The efficacy of the model relies on the caliber and comprehensiveness of the input data, together with the particulars of the problem. | Requires training for new data points |

## 4. PROPOSED METHODOLOGY

In the proposed study, we identify anomalies in data collected from a WSN where multiple nodes collect environmental data. The data from nearby nodes are correlated in space and time. This research focuses on the use of KNN and Z score algorithms. This section describes the complete research process from problem formulation to results on the deployed models with a description of the methods used.

### 4.1. Problem formulation

Consider a specific area where several sensor nodes are located. They transmit information about the environment and communicate with each other via wireless technologies. A hierarchical topology with some number of nodes underpins the WSN under study. Each node $n_i$ is linked to a group of nodes that are adjacent in space. Each sensor node is presumed to be equipped with $p(p >= 2)$ and connected to various types of sensors. These sensors gather $p$-dimensional data at each sampling interval. In a certain area, the collected data from nearby nodes are strongly correlated in both space and time. This correlation applies to qualities such as pressure, humidity, and temperature. At each sampling time $t$, each node $n_i$ has a vector of data $x_{tm}^{i}$. The $l$ adjacent nodes $n_i$ are represented as $a_{ij}$ in space, where $j = 1,2,\dots,l$. At the $k_{th}$ sampling time, $\{x_{tm}^{i}, x_{tm}^{i1}, x_{tm}^{i2}, \dots, x_{tm}^{ij}\}$ denotes the $m$-dimensional data vectors in $\{N_i, N_{i1}, N_{i2}, \dots, N_{ij}\}$. The problem is to determine the normal or non-normal value for each new perceived data vector $x_{tm}^{i}$ node $n_i$.

In the present study, we apply the KNN algorithm and Z score. The KNN algorithm is a nonparametric algorithm based on the idea of observation. Nodes that are close to each other mostly belong to the same class [64]. The Z score algorithm is defined as a measure of the deviation of a variety of experimental observations from the result with the highest probability, the mean [65].

### 4.2. An experimental study

The experimental study was conducted using the open-source RT-IoT2022 dataset. The dataset contains network traffic data that were collected in real time from the interaction of various sensors, covering both normal and abnormal network behavior. The data are a single-formatted two-dimensional dataset in which the columns are features and the rows are examples. In total, the dataset has 4,944 rows×85 columns. Table 3 lists the dataset contains normal data such as ThingSpeak-LED, MQTT-Temp, Amazon-Alexa, and Wipro-Bulb data and non-normal data such as secure shell (SSH) brute-force and DDoS data [66].

Technical specifications of the computer on which the study was conducted:

- CPU: 12th Gen Intel(R) Core™ i7-12700H -2.3 GHz
- RAM: 16 GB
- OS: Windows 11 Pro 64-bit

Table 3. The dataset contains both attack patterns and normal patterns

| Network behaviors | PCAP | Protocol | | Service | | | | | | Patterns |
|---|---|---|---|---|---|---|---|---|---|---|
| | | TCP | UDP | DNS | HTTP | MQTT | SSL | DNS | SSH | |
| SSH brute-force | 1 564 | + | + | + | - | - | - | + | + | Anomaly |
| DDoS | 1 786 | + | - | - | + | - | - | - | - | Anomaly |
| MQTT-Temp | 8 162 | + | - | - | - | + | - | - | - | Normal |
| ThingSpeak-LED | 10 526 | + | + | + | + | - | - | + | - | Normal |
| Amazon-Alexa | 6 056 | + | + | + | + | - | - | + | - | Normal |
| Wipro-Bulb | 1 265 | + | + | + | - | - | + | - | - | Normal |

### 4.3. Proposed system

For practical realization of the task, a model was developed using two methods to detect anomalies. The system shown in Figure 5 processes the dataset and prepares it for further processing. Then, the KNN and Z score algorithms are applied to identify anomalies in the dataset by month.

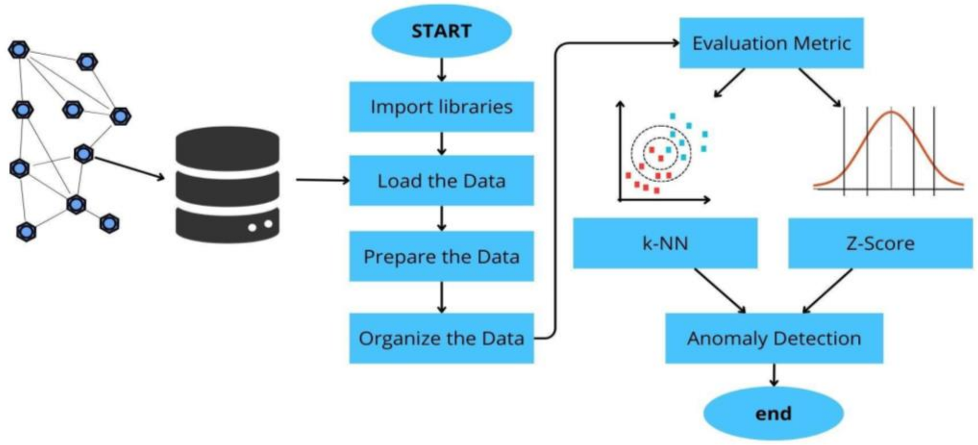
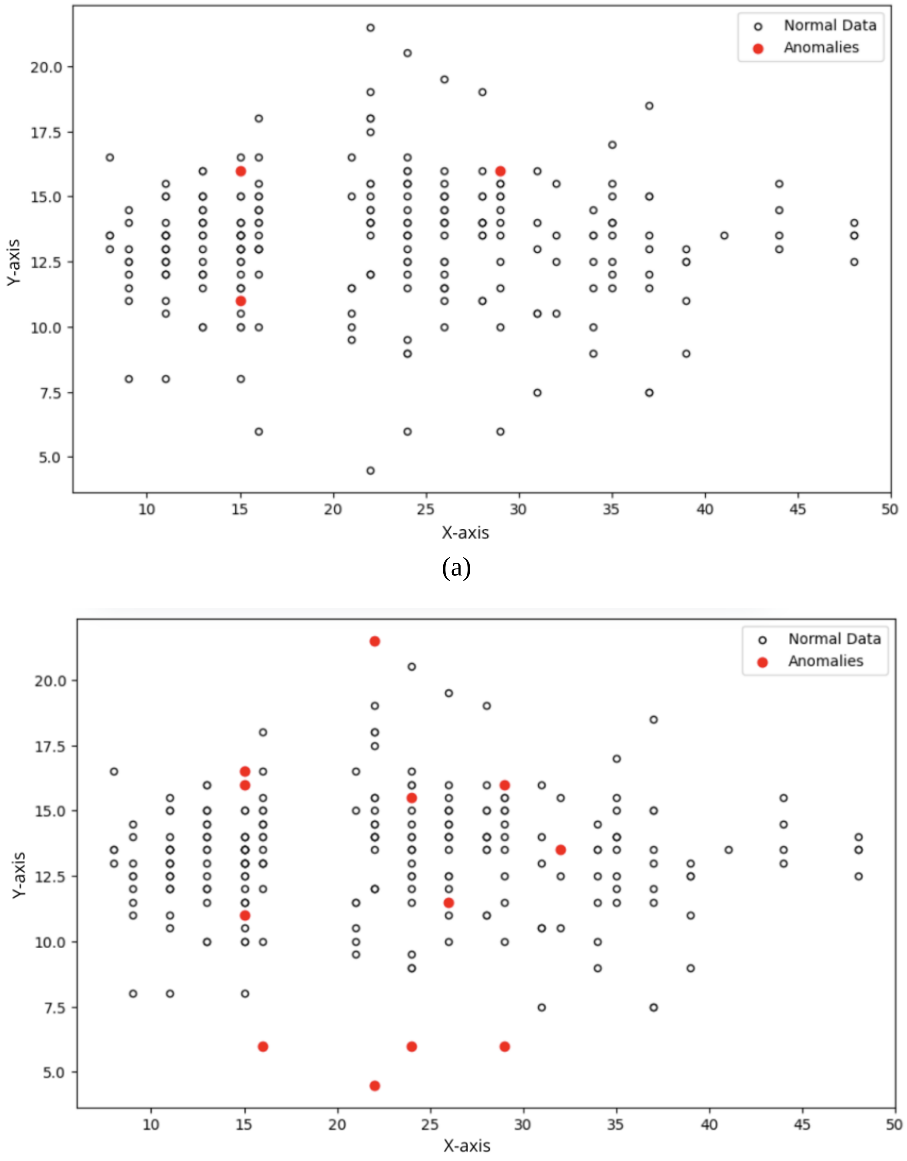


Figure 5. General types of anomaly detection techniques

Figure 6 exhibits the experimental results using two different models. Specifically, Figure 6(a) demonstrates the anomaly detection results achieved with the KNN method. This part of the figure visually represents how the KNN algorithm identified anomalies. In contrast, Figure 6(b) displays the results generated by the Z score method. Therefore, Figure 6 offers a comparison of anomaly detection performance between KNN and Z score approaches.

The visual models show anomalous data points; in the first visual model, anomalous points were detected via the KNN method and in the second plot, the Z score was used. A diverse range of values is demonstrated by the fact that the points are located throughout the area and some outliers are observed where the level of values is much higher than normal. Detecting and removing anomalies improves model accuracy.



(a)

(b)

Figure 6. Experimental results (a) KNN and (b) Z score

## 5. CONCLUSION

Providing security services in WSN via anomaly detection systems is a challenging task. In this work, anomaly detection in WSN was investigated via two algorithms. This study revealed that the Z score has a high anomaly detection rate with respect to the KNN algorithm. The overall accuracy of anomaly detection via KNN was 43.7%, and that via the Z score algorithm was 98.9%. Anomalies can be caused by errors in data collection in the WSN; therefore, evaluating the quality of the data source is important. Anomalies can significantly affect the accuracy of machine learning models. Importantly, to note that models trained on certain datasets may not work well on other datasets. However, removing anomalies will definitely improve the accuracy and generalizability of the model. In summary, to obtain good results, it is necessary to analyze the dataset used carefully after the anomaly removal step.

The present study makes several significant contributions to the field of anomaly detection. We present a new anomaly detection method and describe the various problems that can arise when different anomaly detection methods are used. We have demonstrated an anomaly detection methodology that can be applied to different problems. We believe that the presented research results make valuable contributions to the field of ensuring the security of WSN in terms of anomaly detection. The work was written as part of a scientific project and in the future, we plan to present the results of our research on other methods for detecting anomalies in WSN. The study of existing and developed new algorithms will increase the percentage of attack detection and improve the overall level of WSN security.

## ACKNOWLEDGEMENTS

This research has been funded by the Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan (AP19680345).

## REFERENCES

[1] U. Panahi and C. Bayılmış, “Enabling secure data transmission for wireless sensor networks based IoT applications,” *Ain Shams Engineering Journal*, vol. 14, no. 2, p. 101866, Mar. 2023, doi: 10.1016/j.asej.2022.101866.

[2] J. Amutha, S. Sharma, and J. Nagar, “WSN strategies based on sensors, deployment, sensing models, coverage and energy efficiency: Review, approaches and open issues,” *Wireless Personal Communications*, vol. 111, no. 2, pp. 1089–1115, Oct. 2019, doi: 10.1007/s11277-019-06903-z.

[3] Research and Markets, “Wireless sensor network - global strategic business report,” *Global Industry Analysts, Inc*, 2024. https://www.researchandmarkets.com/reports/5303843/wireless-sensor-network-global-strategic#rela4-5868443 (accessed Jul. 4, 2024).

[4] L. Erhan *et al.*, “Smart anomaly detection in sensor systems: a multi-perspective review,” *Information Fusion*, vol. 67, pp. 64–79, Mar. 2021, doi: 10.1016/j.inffus.2020.10.001.

[5] M. S. BenSaleh, R. Saida, Y. H. Kacem, and M. Abid, “Wireless sensor network design methodologies: a survey,” *Journal of Sensors*, vol. 2020, pp. 1–13, Jan. 2020, doi: 10.1155/2020/9592836.

[6] M. Keerthika and D. Shanmugapriya, “Wireless sensor networks: active and passive attacks - vulnerabilities and countermeasures,” *Global Transitions Proceedings*, vol. 2, no. 2, pp. 362–367, Nov. 2021, doi: 10.1016/j.gltp.2021.08.045.

[7] D. Kandris, C. Nakas, D. Vomvas, and G. Koulouras, “Applications of wireless sensor networks: an up-to-date survey,” *Applied System Innovation*, vol. 3, no. 1, Feb. 2020, doi: 10.3390/asi3010014.

[8] Anitha C L and R. Sumathi, “Anomaly detection in WSN IoT (internet of things) environment through a consensus-based anomaly detection approach,” *Multimedia Tools and Applications*, vol. 83, no. 20, pp. 58915–58934, Dec. 2023, doi: 10.1007/s11042-023-17894-2.

[9] S. M. S. Bukhari, M. H. Zafar, M. A. Houran, Z. Qadir, S. K. R. Moosavi, and F. Sanfilippo, “Enhancing cybersecurity in Edge IIoT networks: an asynchronous federated learning approach with a deep hybrid detection model,” *Internet of Things*, vol. 27, p. 101252, Oct. 2024, doi: 10.1016/j.iot.2024.101252.

[10] J. Xavier A, S. Vanitha N, Sudha G, and Birunda M, “Combined localization and clustering approach for reduced energy presumption in heterogeneous IoT,” *Physica Scripta*, vol. 99, no. 7, p. 75222, Jun. 2024, doi: 10.1088/1402-4896/ad4f2b.

[11] T. Bapu B R, M. Shankar, M. P. Aravinth, T. Eashwar, S. Subash, and B. Sethu, “An efficient network anomaly detection scheme over heavy traffic wireless sensor network environment,” in *2023 9th International Conference on Smart Structures and Systems (ICSSS)*, Nov. 2023, pp. 1–8, doi: 10.1109/icsss58085.2023.10407897.

[12] S. Gayathri and D. Surendran, “Unified ensemble federated learning with cloud computing for online anomaly detection in energy-efficient wireless sensor networks,” *Journal of Cloud Computing*, vol. 13, no. 1, Feb. 2024, doi: 10.1186/s13677-024-00595-y.

[13] A. R. Affane M. and H. Satori, “Machine learning attack detection based-on stochastic classifier methods for enhancing of routing security in wireless sensor networks,” *Ad Hoc Networks*, vol. 163, p. 103581, Oct. 2024, doi: 10.1016/j.adhoc.2024.103581.

[14] M. Matar, T. Xia, K. Huguenard, D. Huston, and S. Wshah, “Multi-head attention based Bi-LSTM for anomaly detection in multivariate time-series of WSN,” in *2023 IEEE 5th International Conference on Artificial Intelligence Circuits and Systems (AICAS)*, Jun. 2023, pp. 1–5, doi: 10.1109/aicas57966.2023.10168670.

[15] M. Safaei *et al.*, “A systematic literature review on outlier detection in wireless sensor networks,” *Symmetry*, vol. 12, no. 3, p. 328, Feb. 2020, doi: 10.3390/sym12030328.

[16] 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, Jan. 2022, doi: 10.3390/jsan11010004.

[17] S. Mohan, A. Manke, S. Verma, and K. Baskar, “Machine learning at the edge: GANs for anomaly detection in wireless sensor networks,” in *Enhancing Security in Public Spaces Through Generative Adversarial Networks (GANs)*, IGI Global, 2024, pp. 305–317.

[18] S. Bacha, A. Aljuhani, K. Ben Abdellafou, O. Taouali, N. Liouane, and M. Alazab, “Anomaly-based intrusion detection system in IoT using kernel extreme learning machine,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 15, no. 1, pp. 231–242, May 2022, doi: 10.1007/s12652-022-03887-w.

[19] M. P. S. Bhatia and S. R. Sangwan, “Soft computing for anomaly detection and prediction to mitigate IoT-based real-time abuse,” *Personal and Ubiquitous Computing*, vol. 28, no. 1, pp. 123–133, May 2021, doi: 10.1007/s00779-021-01567-8.

[20] M. M. Inuwa and R. Das, “A comparative analysis of various machine learning methods for anomaly detection in cyber attacks on IoT networks,” *Internet of Things*, vol. 26, p. 101162, Jul. 2024, doi: 10.1016/j.iot.2024.101162.

[21] M. Dissem, M. Amayri, and N. Bouguila, "Neural architecture search for anomaly detection in time-series data of smart buildings: A reinforcement learning approach for optimal autoencoder design," *IEEE Internet of Things Journal*, vol. 11, no. 10, pp. 18059–18073, May 2024, doi: 10.1109/jiot.2024.3360882.

[22] A. R. Affane M., H. Satori, Y. Boutazart, A. Ezzine, and K. Satori, "Machine learning-based attack detection for wireless sensor network security using hidden Markov models," *Wireless Personal Communications*, vol. 135, no. 4, pp. 1965–1992, Apr. 2024, doi: 10.1007/s11277-024-10999-3.

[23] A. Algarni, T. Acarer, and Z. Ahmad, "An edge computing-based preventive framework with machine learning integration for anomaly detection and risk management in maritime wireless communications," *IEEE Access*, vol. 12, pp. 53646–53663, 2024, doi: 10.1109/access.2024.3387529.

[24] S. Alangari, "An unsupervised machine learning algorithm for attack and anomaly detection in IoT sensors," *Wireless Personal Communications*, Feb. 2024, doi: 10.1007/s11277-023-10811-8.

[25] V. Sivagaminathan, M. Sharma, and S. K. Henge, "Intrusion detection systems for wireless sensor networks using computational intelligence techniques," *Cybersecurity*, vol. 6, no. 1, Oct. 2023, doi: 10.1186/s42400-023-00161-0.

[26] A. R. Moundounga, H. Satori, Y. Boutazart, and E. Abderrahim, "Malicious attack detection based on continuous hidden Markov models in wireless sensor networks," *Microprocessors and Microsystems*, vol. 101, p. 104888, Sep. 2023, doi: 10.1016/j.micpro.2023.104888.

[27] 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, Mar. 2023, doi: 10.1007/s11277-023-10253-2.

[28] 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.

[29] S. Scanzio *et al.*, "Wireless sensor networks and TSCH: a compromise between reliability, power consumption, and latency," *IEEE Access*, vol. 8, pp. 167042–167058, 2020, doi: 10.1109/access.2020.3022434.

[30] A. Adamova, T. Zhukabayeva, and Y. Mardenov, "Machine learning in action: an analysis of its application for fault detection in wireless sensor networks," in *2023 IEEE International Conference on Smart Information Systems and Technologies (SIST)*, May 2023, pp. 506–511, doi: 10.1109/sist58284.2023.10223548.

[31] Y. Mardenov, A. Adamova, T. Zhukabayeva, and M. Othman, "Enhancing fault detection in wireless sensor networks through support vector machines: a comprehensive study," *Journal of Robotics and Control (JRC)*, vol. 4, no. 6, pp. 868–877, Dec. 2023, doi: 10.18196/jrc.v4i6.20216.

[32] A. Gaddam, T. Wilkin, M. Angelova, and J. Gaddam, "Detecting sensor faults, anomalies and outliers in the Internet of Things: A survey on the challenges and solutions," *Electronics*, vol. 9, no. 3, p. 511, Mar. 2020, doi: 10.3390/electronics9030511.

[33] A. Castellani, S. Schmitt, and S. Squartini, "Real-world anomaly detection by using digital twin systems and weakly supervised learning," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 7, pp. 4733–4742, Jul. 2021, doi: 10.1109/tii.2020.3019788.

[34] A. B. Nassif, M. A. Talib, Q. Nasir, and F. M. Dakalbab, "Machine learning for anomaly detection: a systematic review," *IEEE Access*, vol. 9, pp. 78658–78700, 2021, doi: 10.1109/access.2021.3083060.

[35] W. Khan and M. Haroon, "An unsupervised deep learning ensemble model for anomaly detection in static attributed social networks," *International Journal of Cognitive Computing in Engineering*, vol. 3, pp. 153–160, Jun. 2022, doi: 10.1016/j.ijcce.2022.08.002.

[36] Y. Lu *et al.*, "Semi-supervised machine learning aided anomaly detection method in cellular networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 8, pp. 8459–8467, Aug. 2020, doi: 10.1109/tvt.2020.2995160.

[37] A. Chirayil, R. Maharjan, and C.-S. Wu, "Survey on anomaly detection in wireless sensor networks (WSNs)," in *2019 20th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)*, Jul. 2019, pp. 150–157, doi: 10.1109/snpd.2019.8935827.

[38] Y. Sun and Y. Chen, "Detection of wormhole attacks in wireless sensor networks based on anomaly detection algorithms," in *2022 2nd International Conference on Consumer Electronics and Computer Engineering (ICCECE)*, Jan. 2022, pp. 777–782, doi: 10.1109/iccece54139.2022.9712659.

[39] A. Adamova, T. Zhukabayeva, Z. Mukanova, and Z. Oralbekova, "Enhancing internet of things security against structured query language injection and brute force attacks through federated learning," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 15, no. 1, pp. 1187–1199, Feb. 2025, doi: 10.11591/ijece.v15i1.pp1187-1199.

[40] P. Biswas and T. Samanta, "Anomaly detection using ensemble random forest in wireless sensor network," *International Journal of Information Technology*, vol. 13, no. 5, pp. 2043–2052, Jun. 2021, doi: 10.1007/s41870-021-00717-8.

[41] M. Karthikeyan, D. Manimegalai, and K. RajaGopal, "Firefly algorithm based WSN-IoT security enhancement with machine learning for intrusion detection," *Scientific Reports*, vol. 14, no. 1, Jan. 2024, doi: 10.1038/s41598-023-50554-x.

[42] F. M. Ghamry, G. M. El-Banby, A. S. El-Fishawy, F. E. A. El-Samie, and M. I. Dessouky, "A survey of anomaly detection techniques," *Journal of Optics*, vol. 53, no. 2, pp. 756–774, Feb. 2024, doi: 10.1007/s12596-023-01147-4.

[43] A. Abid, S. El Khediri, and A. Kachouri, "Improved approaches for density-based outlier detection in wireless sensor networks," *Computing*, vol. 103, no. 10, pp. 2275–2292, Apr. 2021, doi: 10.1007/s00607-021-00939-5.

[44] D. Samariya and A. Thakkar, "A comprehensive survey of anomaly detection algorithms," *Annals of Data Science*, Nov. 2021, doi: 10.1007/s40745-021-00362-9.

[45] 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.

[46] Y. Chen and S. Li, "A lightweight anomaly detection method based on SVDD for wireless sensor networks," *Wireless Personal Communications*, vol. 105, no. 4, pp. 1235–1256, Feb. 2019, doi: 10.1007/s11277-019-06143-1.

[47] Y. Shi and H. Shen, "Unsupervised anomaly detection for network traffic using artificial immune network," *Neural Computing and Applications*, vol. 34, no. 15, pp. 13007–13027, Mar. 2022, doi: 10.1007/s00521-022-07156-x.

[48] R. SaiSindhuTheja and G. K. Shyam, "An efficient metaheuristic algorithm based feature selection and recurrent neural network for DoS attack detection in cloud computing environment," *Applied Soft Computing*, vol. 100, p. 106997, Mar. 2021, doi: 10.1016/j.asoc.2020.106997.

[49] I. Al-Turaiki and N. Altwaijry, "A convolutional neural network for improved anomaly-based network intrusion detection," *Big Data*, vol. 9, no. 3, pp. 233–252, Jun. 2021, doi: 10.1089/big.2020.0263.

[50] C. Yao, Y. Yang, K. Yin, and J. Yang, "Traffic anomaly detection in wireless sensor networks based on principal component analysis and deep convolution neural network," *IEEE Access*, vol. 10, pp. 103136–103149, 2022, doi: 10.1109/access.2022.3210189.

[51] M. Jain, G. Kaur, and V. Saxena, "A k-means clustering and SVM based hybrid concept drift detection technique for network anomaly detection," *Expert Systems with Applications*, vol. 193, p. 116510, May 2022, doi: 10.1016/j.eswa.2022.116510.

[52] D. Mishra, P. Roy, and A. Mishra, "Trap-based anomaly detection mechanism for wireless sensor network," *ASEAN Engineering Journal*, vol. 14, no. 2, pp. 167–172, May 2024, doi: 10.11113/aej.v14.20997.

[53] J. Kim, Y. Moon, and H. Ko, "Correlation-based advanced feature analysis for wireless sensor networks," *The Journal of Supercomputing*, vol. 80, no. 7, pp. 9812–9828, Dec. 2023, doi: 10.1007/s11227-023-05739-6.

[54] B. Chander and G. Kumaravelan, "Outlier detection strategies for WSNs: A survey," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 8, pp. 5684–5707, Sep. 2022, doi: 10.1016/j.jksuci.2021.02.012.

[55] Z. Wei, M. Li, and J. Wei, "Long-short term anomaly detection in wireless sensor networks based on spatio-temporal correlation in IoT systems," in *2023 6th International Conference on Artificial Intelligence and Pattern Recognition (AIPR)*, Sep. 2023, pp. 1042–1048, doi: 10.1145/3641584.3641739.

[56] U. Gupta, V. Bhattacharjee, and P. S. Bishnu, "Outlier detection in wireless sensor networks based on neighbourhood," *Wireless Personal Communications*, vol. 116, no. 1, pp. 443–454, Aug. 2020, doi: 10.1007/s11277-020-07722-3.

[57] S. Otoum, B. Kantarci, and H. Mouftah, "A comparative study of AI-based intrusion detection techniques in critical infrastructures," *ACM Transactions on Internet Technology*, vol. 21, no. 4, pp. 1–22, Jul. 2021, doi: 10.1145/3406093.

[58] S. Ifzarne, H. Tabbaa, I. Hafidi, and N. Lamghari, "Anomaly detection using machine learning techniques in wireless sensor networkss," *Journal of Physics: Conference Series*, vol. 1743, no. 1, p. 12021, Jan. 2021, doi: 10.1088/1742-6596/1743/1/012021.

[59] D. Widhalm, K. M. Goeschka, and W. Kastner, "SoK: a taxonomy for anomaly detection in wireless sensor networks focused on node-level techniques," Aug. 2020, doi: 10.1145/3407023.3407027.

[60] S. Bajpai, P. Krishna Murthy, and N. Kumar, "AnomGraphAdv: Enhancing anomaly and network intrusion detection in wireless networks using adversarial training and temporal graph networks," in *Proceedings of the 17th ACM Conference on Security and Privacy in Wireless and Mobile Networks*, May 2024, pp. 113–122, doi: 10.1145/3643833.3656130.

[61] A. M. Jubair *et al.*, "Optimization of clustering in wireless sensor networks: Techniques and protocols," *Applied Sciences*, vol. 11, no. 23, p. 11448, Dec. 2021, doi: 10.3390/app112311448.

[62] M. Al Samara, I. Bennis, A. Abouaissa, and P. Lorenz, "An efficient outlier detection and classification clustering-based approach for WSN," in *2021 IEEE Global Communications Conference (GLOBECOM)*, Dec. 2021, pp. 1–6, doi: 10.1109/globecom46510.2021.9685756.

[63] M. Al Samara, I. Bennis, A. Abouaissa, and P. Lorenz, "Sa-o2dca: Seasonal adapted online outlier detection and classification approach for WSN," *Journal of Network and Systems Management*, vol. 32, no. 2, Mar. 2024, doi: 10.1007/s10922-024-09801-3.

[64] G. Liu, H. Zhao, F. Fan, G. Liu, Q. Xu, and S. Nazir, "An enhanced intrusion detection model based on improved kNN in WSNs," *Sensors*, vol. 22, no. 4, p. 1407, Feb. 2022, doi: 10.3390/s22041407.

[65] A. S. Yaro, F. Maly, and P. Prazak, "Outlier detection in time-series receive signal strength observation using z-score method with Sn scale estimator for indoor localization," *Applied Sciences*, vol. 13, no. 6, p. 3900, Mar. 2023, doi: 10.3390/app13063900.

[66] B. S. Sharmila and R. Nagapadma, "Quantized autoencoder (QAE) intrusion detection system for anomaly detection in resource-constrained IoT devices using RT-IoT2022 dataset," *Cybersecurity*, vol. 6, no. 1, Sep. 2023, doi: 10.1186/s42400-023-00178-5.

## BIOGRAPHIES OF AUTHORS

**Tamara Zhukabayeva** received the Ph.D. degree from Satbayev University, Kazakhstan. She is currently an associate professor in informatics, computer engineering and control, 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@gmail.com.

**Aigul Adamova** received the M.S. degree in informatics from L. Gumilyov Eurasian National University, Kazakhstan, in 2006 and the Ph.D. degrees in computing and software from L.Gumilyov Eurasian National University, Kazakhstan, in 2016. Currently, she is a postdoctoral researcher and an assistant professor at Department of Computer Engineering, Astana IT University, Kazakhstan. She has about 40 published papers in refereed journals and conferences. She served as a reviewer for international conferences, including IEEE: SIST 2023, SIST 2024. Her research areas are information security of internet of things, wireless sensor network, embedded system, cyberphysical system and computer vision. She can be contacted at email: aigul.adamova@astanait.edu.kz.

**Lyazzat Zholshiyeva** in 2003, she graduated from the Taraz State University named after M. Kh. Dulati with a degree in mathematics and computer sciences. In 2012, she received a master's degree in mechanical engineering. In 2020, she graduated from the doctoral program Astana International University, specialty 8D06101- "Computing and software". Her research interests include computer vision, machine learning, IoT. Researcher at the International Science Complex "Astana". She can be contacted at email: lazzat.zhol.81@gmai.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.

**Nurdaulet Karabayev** Junior researcher at the International Science Complex “Astana.” His research focuses on developing efficient information retrieval systems and enhancing security in wireless sensor networks. He has contributed to the creation of a full-text retrieval system for digital libraries, aiming to improve access to relevant information. He can be contacted at email: 222240@astanait.edu.kz.

**Dilaram Baumuratova** Ph.D., senior lecturer at the Pedagogical Institute of Astana International University, Kazakhstan. Her research focuses on the integration of cloud computing technologies in technical and vocational education systems. Her work explores the theoretical and practical foundations of utilizing cloud solutions to enhance teaching methodologies and curriculum content in professional education. She can be contacted at email: baumuratova.d@gmail.com.