JOURNALS | Journal of Smart Sensors and Computing

Volume 2 Issue 4 (December 2026)
J. Smart Sens. Comput., 2026, 2(4), 26219
DOI : https://doi.org/10.64189/ssc.26219

Research Article

OPEN ACCESS

An Energy-Aware Adaptive Sampling and Anomaly Detection Machine Learning Framework for Intelligent Wireless Sensor Networks

Olarewaju Peter Ayeoribe[1,2,*]

[1] *Department of Electrical and Electronics Engineering, Federal University Oye-Ekiti, 371104, Nigeria*
[2] *Transmitter Laboratory, Peters A.O. Broadcasting Company Ltd, Nigeria*

*Corresponding Author
Olarewaju Peter Ayeoribe
ayeoribe.olarewaju@fuoye.edu.ng

Received: 10 October 2026
Revised: 10 November 2026
Accepted: 17 November 2026
Published Online: 18 November 2026

Citation

O. P. Ayeoribe, An energy-aware adaptive sampling and anomaly detection machine learning framework for intelligent wireless sensor networks, *Journal of Smart Sensors and Computing*, 2026, 2(4), 26219, https://doi.org/10.64189/ssc.26219.

Open Access

This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits the non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as appropriate credit to the original author(s) and the source is given by providing a link to the Creative Commons License and changes need to be indicated if there are any. The images or other third-party material in this article are included in the article's Creative Commons License, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons License and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this License, visit: https://creativecommons.org/licenses/by-nc/4.0/

© The Author(s) 2026

Abstract

Energy efficiency and reliable anomaly detection remain important challenges in resource-constrained wireless sensor networks (WSNs), where continuous data acquisition and abnormal network traffic can increase communication and processing demands. This study presents an energy-aware adaptive sampling and machine learning-based anomaly detection framework for intelligent WSNs. The proposed framework integrates condition-dependent sampling with Random Forest classification to reduce unnecessary data acquisition while supporting the identification of abnormal network traffic. The anomaly detection component was implemented using the WSN-DS dataset, comprising 374,661 records generated through NS-2 simulation using the Low-Energy Adaptive Clustering Hierarchy (LEACH) protocol. The dataset contained five traffic categories: normal, blackhole, flooding, grayhole, and scheduling (TDMA). A Random Forest classifier comprising 300 decision trees was configured using Gini impurity, square-root feature selection, and balanced subsampling to address class imbalance. Adaptive sampling performance was evaluated under stable, moderately varying, and highly dynamic operating conditions. The results showed sampling reductions of 60%, 40%, and 15%, respectively. Across the three equally weighted conditions, adaptive sampling reduced the total number of samples from 30,000 to 18,500, representing an aggregate reduction of 38.33% compared with fixed-rate sampling. These findings demonstrate the potential of condition-dependent sampling to improve sampling efficiency while retaining greater data acquisition activity under dynamic conditions. Further classification and network-level evaluation is required to quantify the corresponding effects on detection performance, energy consumption, packet transmission, and network lifetime.

Keywords: Adaptive sampling; Anomaly detection; Energy efficiency; Machine learning; Wireless sensor networks.

1. Introduction

1.1 Background

Wireless sensor networks (WSNs) constitute an important component of intelligent monitoring and Internet of Things (IoT) infrastructures, supporting applications such as environmental surveillance, industrial automation, smart-city management, healthcare monitoring and water-quality assessment. Through distributed sensor nodes, WSNs facilitate the continuous acquisition, processing and transmission of information from physical environments. However, their operational effectiveness is constrained by limited battery capacity, restricted processing resources, communication bandwidth and variable network conditions. These limitations make energy-efficient sensing, reliable data acquisition and timely anomaly detection essential requirements for sustainable WSN operation.[1–3]

Energy consumption in WSNs is strongly influenced by sensing frequency, local computation and wireless communication. Continuous data acquisition at predetermined sampling intervals may generate redundant observations during relatively stable operating periods, resulting in unnecessary processing and communication overhead. Conversely, excessively reducing the sampling frequency may prevent the timely identification of short-duration abnormalities and rapidly changing environmental conditions. Consequently, selecting an appropriate sampling frequency represents an important resource-management problem in applications where sensor nodes are expected to operate for extended periods.

Security and operational reliability introduce additional challenges. WSNs may experience denial-of-service attacks, abnormal traffic patterns, compromised nodes, routing-related attacks and sensor faults. Machine learning (ML) techniques have therefore received considerable attention for identifying abnormal behaviour through the analysis of sensor measurements and network characteristics. Existing studies have investigated ML-based WSN security,[1] AI-driven anomaly detection in sensor and IoT environments,[2] and broader machine-learning approaches to WSN security.[3] Further investigations have examined lightweight detection techniques, ensemble-based attack identification and energy-aware learning mechanisms.[4-9] These studies demonstrate the potential of intelligent detection techniques for improving network reliability.

Nevertheless, the computational efficiency of an anomaly detection algorithm does not necessarily address the energy consumed during unnecessary sensor measurements and data transmissions. A detection model may operate efficiently while continuing to receive observations at an unnecessarily high sampling frequency. Adaptive sampling provides an alternative approach by adjusting data acquisition according to changes in sensor measurements, available energy and monitoring requirements. Under relatively stable conditions, sampling activity can be reduced, whereas increased measurement frequency can be activated when significant variations or potential abnormal events are identified.

Integrating adaptive sampling with ML-based anomaly detection provides an opportunity to coordinate data acquisition and abnormality identification within a common monitoring architecture. Such integration requires a decision-making mechanism that considers observed signal variations, anomaly-related information and residual energy when selecting subsequent sampling intervals. The mechanism must also account for the possibility that reducing the sampling frequency may delay or prevent the detection of transient abnormalities. Accordingly, this study proposes an energy-aware adaptive sampling and ML-based anomaly detection framework for intelligent WSNs. The framework considers sampling frequency as an integral component of anomaly monitoring and resource management rather than treating sensing efficiency and detection performance as independent optimisation problems. Its evaluation focuses on the relationship between energy consumption, communication overhead, anomaly detection reliability, detection latency and network lifetime.

1.2 Literature review

Recent studies have investigated ML-based approaches for strengthening the security, reliability, and operational efficiency of WSNs and IoT infrastructures. A lightweight ML approach has been developed for detecting denial-of-service attacks in WSNs.[4] Cluster-based detection using variable-selection ensemble learning has also been investigated to improve the identification of DoS attacks.[5] Furthermore, enhanced extreme learning machine techniques have been explored for detecting vampire attacks, which can drain network energy through malicious routing behaviour.[6] These investigations demonstrate the importance of selecting suitable learning algorithms for different threat categories. However, their primary emphasis on attack detection does not, by itself, establish an integrated mechanism for dynamically controlling the sampling process according to energy availability and the changing likelihood of abnormal events.

Energy efficiency has also received attention in the development of intelligent anomaly detection systems. An energy-aware and explainable automated ML approach has been investigated for anomaly detection in healthcare IoT environments.[7] Federated ensemble learning combined with cloud computing has been explored for online anomaly detection in energy-efficient

WSNs.[8] In addition, an adaptive deep learning framework has been proposed for energy-efficient anomaly detection in IoT sensor networks.[9] These contributions indicate that energy considerations are increasingly being incorporated into intelligent monitoring architectures. Nevertheless, energy-aware learning, distributed model execution, and adaptive data acquisition represent distinct design dimensions. Reducing the computational cost of model execution or distributing learning operations does not necessarily reduce the energy consumed by unnecessary sensing and data transmission. This distinction motivates the investigation of sampling policies that directly respond to network conditions and detection requirements.

Other research has examined ML-enabled monitoring and anomaly detection in broader communication and sensing environments. An edge computing-based preventive framework integrating ML has been proposed for anomaly detection and risk management in maritime wireless communications.[10] ML-based methods have also been investigated for anomaly detection in 6G networks,[11] while ML applications in WSNs supporting smart-city services have been reviewed.[12] These studies demonstrate the applicability of intelligent analysis across diverse network architectures and application domains. However, there remains a need for a WSN-specific framework that coordinates sampling frequency, energy expenditure, and anomaly detection within a common decision-making process.

The literature further demonstrates the importance of improving detection robustness and the interpretability of intelligent monitoring systems. A comprehensive security framework based on adaptive ML has been investigated for low-overhead anomaly detection in residential networks.[13] A reproducible neuro-fuzzy framework has also been reported for interpretable anomaly detection in WSN data streams.[14] These approaches are relevant to the development of resource-conscious and understandable monitoring systems. However, the extent to which adaptive detection mechanisms can be combined with event-responsive sampling and explicit energy-management objectives requires further investigation.

Security mechanisms based on complementary technologies have also been examined. ML and blockchain have been reviewed as approaches for securing WSNs,[15] while an ML-based intrusion detection system has been investigated for WSNs.[16] An intelligent intrusion detection framework has also been reported.[17] Furthermore, a lightweight ML-based anomaly detection and attack classification framework has been proposed for enhancing WSN security.[18] Collectively, these studies reinforce the importance of detecting malicious behavior and improving network security. Nevertheless, intrusion detection and anomaly detection may differ in their objectives: intrusion detection typically focuses on identifying security violations, whereas anomaly detection can additionally identify unexpected system behavior, sensor faults, and unusual environmental observations. A comprehensive intelligent monitoring architecture should therefore accommodate broader abnormality detection while controlling the energy cost of collecting the data required for analysis.

Beyond security, WSNs support important environmental and infrastructure-monitoring applications. A heterogeneous wireless multimedia sensor network framework has been developed for forest-fire detection,[19] demonstrating the importance of timely observations in event-sensitive monitoring. IoT-enabled WSNs have also been investigated for smart water quality monitoring.[20] These applications illustrate why fixed-rate sampling can be inefficient in some conditions while insufficient in others: stable periods may not require frequent observations, whereas rapidly changing environmental conditions may require increased sampling frequency. Accordingly, adaptive sampling can potentially reduce redundant measurements while preserving the observations necessary for reliable detection.

Additional research has explored optimized learning-based anomaly recognition for WSN security[21-22] and lightweight ML-based anomaly detection and attack classification.[18] Other contributions have examined energy-aware learning,[7] adaptive deep learning,[9] and interpretable anomaly detection.[14] Although these approaches provide useful foundations for intelligent monitoring, their differing objectives and evaluation settings indicate the need for a unified framework that explicitly coordinates energy-aware sampling and anomaly detection.

1.3 Research gap

Despite advances in ML-based anomaly detection and energy-efficient WSN management, important challenges remain in coordinating data acquisition with anomaly detection and energy constraints.

First, existing detection-oriented approaches primarily concentrate on identifying abnormal sensor behaviour, network intrusions or malicious traffic.[4–6,15–18] Although these techniques improve the ability to recognise abnormal events, their detection performance does not necessarily reduce the energy consumed by continuous sensing and communication.

Second, energy-aware learning approaches have investigated computational optimisation, distributed processing and efficient model execution.[7–9] However, reducing the computational requirements of an ML model does not automatically minimise redundant sensor observations or the associated transmission overhead. Sampling frequency therefore remains an

important design variable in resource-constrained monitoring systems.

Third, fixed-rate sampling does not adequately respond to variations in environmental conditions or the changing likelihood of abnormal events. High sampling frequencies may waste energy during stable operating periods, whereas low sampling frequencies may increase detection latency and the probability of missing transient abnormalities.

Finally, evaluating energy efficiency independently of anomaly detection performance provides an incomplete assessment of WSN sustainability. A meaningful framework must determine whether reductions in sensing and communication costs can be achieved while maintaining acceptable detection reliability and response time.

The research gap addressed in this study is therefore the need for a coordinated monitoring framework in which adaptive sampling, anomaly detection information and residual-energy conditions jointly influence data acquisition decisions. The proposed framework investigates this relationship through comparative evaluation of sampling efficiency, energy consumption, detection performance, communication overhead and network lifetime.

1.4 Aim and objectives

The aim of this study is to develop an energy-aware adaptive sampling and machine-learning-based anomaly detection framework for intelligent WSNs that reduces unnecessary sensing and communication activities while maintaining reliable anomaly detection under changing network conditions.

The specific objectives are to:

1. Develop an adaptive sampling mechanism that adjusts sensor data acquisition according to signal variations, anomaly indicators and residual energy.
2. Integrate a specified ML-based anomaly detection algorithm for identifying abnormal sensor observations and potentially suspicious network behaviour.

Establish a coordinated decision-making mechanism that balances sampling frequency, energy consumption, detection reliability and response time.

3. Evaluate the proposed framework against clearly defined baseline configurations under equivalent operating conditions.
4. Quantitatively assess energy consumption, sampling reduction, packet transmissions, precision, recall, F1-score, accuracy, detection latency and network lifetime.
5. Investigate the trade-offs between energy efficiency and detection reliability through statistical, ablation and sensitivity analyses.

1.5 Research novelty and contributions

The specific contribution of this research is the formulation of a coordinated decision-making architecture in which adaptive sensor sampling is informed by both anomaly-related information and residual-energy conditions. Rather than considering data acquisition, anomaly detection and energy management as separate operations, the framework links their decisions through a feedback mechanism.

The intended technical contributions are:

1. Energy-aware adaptive sampling: A sampling-control mechanism that adjusts data acquisition according to observed signal variability and available sensor-node energy.
2. Anomaly-informed sampling control: Integration of anomaly-related detection information into subsequent sampling decisions, allowing monitoring activity to respond to potential abnormal conditions.
3. Joint resource and detection assessment: A unified evaluation approach that considers sensing and communication energy alongside anomaly detection reliability and response latency.
4. Comparative framework evaluation: Assessment against fixed-rate and partially integrated configurations to identify the contribution of individual framework components.
5. Performance trade-off analysis: Investigation of the relationship between sampling reduction, missed anomalies, detection latency, energy consumption and network lifetime.

The novelty claim is limited to the specific integration and decision-making mechanism implemented in this study. The contribution should not be interpreted as claiming that adaptive sampling or ML-based anomaly detection is independently new.

2. Materials and methods

2.1 Research design and methodological framework

The study adopted a computational framework integrating energy-aware adaptive sampling and supervised machine learning-based anomaly detection for intelligent wireless sensor networks (WSNs). The methodology combined dataset preparation, feature preprocessing, Random Forest classification, adaptive sampling control, energy estimation, and comparative performance evaluation. The principal objective was to coordinate data acquisition and anomaly detection so that unnecessary sensing and communication activities could be reduced without compromising the identification of abnormal network behavior.

Fig. 1 illustrates the proposed methodological framework, showing the sequential integration of dataset preparation, feature preprocessing, Random Forest-based anomaly classification, adaptive sampling control, energy and communication assessment, and comparative performance evaluation.

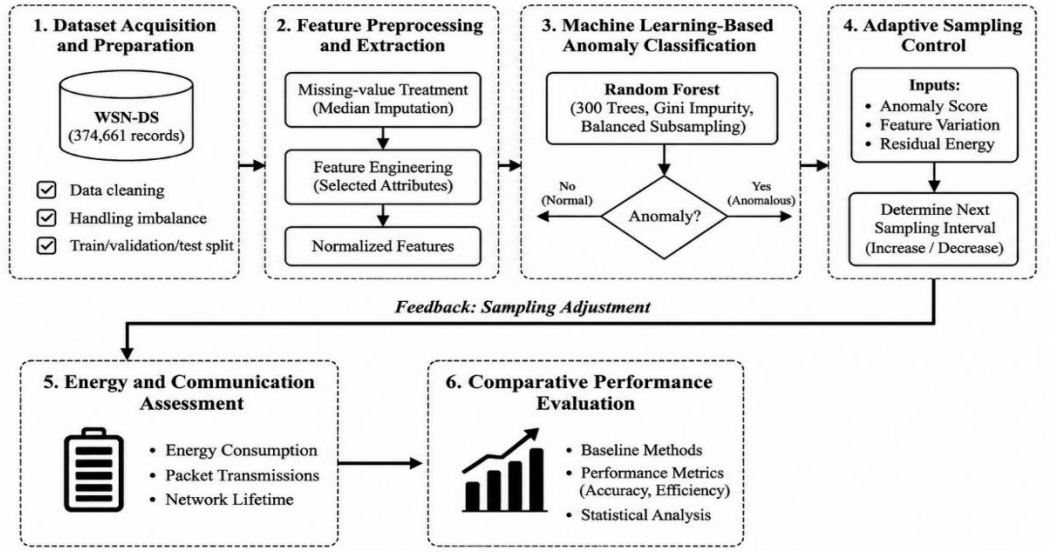


Fig. 1: Methodological block diagram of the proposed energy-aware adaptive sampling and machine learning-based anomaly detection framework for WSNs.

The proposed framework consisted of six interconnected functional stages: (i) dataset acquisition and preparation, (ii) feature preprocessing and extraction, (iii) machine learning-based anomaly classification, (iv) adaptive sampling control, (v) energy and communication assessment, and (vi) comparative performance evaluation. The Random Forest classifier identified normal and anomalous network traffic, while the adaptive controller used the resulting anomaly score, observed feature variation, and available residual energy to determine the subsequent sampling interval.

The overall processing sequence was represented as

$$\mathcal{F}: \mathcal{D} \rightarrow \mathcal{P} \rightarrow \mathcal{X} \rightarrow \mathcal{M} \rightarrow \mathcal{A} \rightarrow \mathcal{E}, \quad (1)$$

where $\mathcal{D}$ denotes the input dataset, $\mathcal{P}$ represents data preprocessing, $\mathcal{X}$ denotes feature extraction, $\mathcal{M}$ represents machine learning-based classification, $\mathcal{A}$ denotes adaptive sampling control, and $\mathcal{E}$ represents the evaluation of energy consumption, detection performance, and communication efficiency.

Unlike conventional fixed-rate monitoring, the proposed framework treated the sampling interval as a dynamic control variable. The sampling controller responded to changes in the monitored data and the anomaly detection output while considering the remaining energy available to the sensor node. Consequently, the framework established a feedback relationship between data acquisition, anomaly classification, and energy management.

2.2 Dataset description and preparation

The Wireless Sensor Network Dataset (WSN-DS) was selected as the public benchmark dataset for evaluating the proposed anomaly detection framework. The dataset contains 374,661 records generated through network simulation using the Network Simulator 2 (NS-2) and the Low-Energy Adaptive Clustering Hierarchy (LEACH) protocol. Five traffic categories are represented: normal traffic, Blackhole, Flooding, Grayhole, and Scheduling (TDMA) attacks.

The class distribution of the dataset is presented in Table 1. Normal traffic accounts for 340,066 records, representing 90.766% of the observations. The remaining records comprise 10,049 Blackhole instances (2.682%), 3,312 Flooding instances (0.884%), 14,596 Grayhole instances (3.896%), and 6,638 Scheduling (TDMA) attack instances (1.772%). The distribution demonstrates substantial class imbalance, with normal traffic considerably outnumbering each individual attack category.

Table 1: Class distribution of the WSN-DS dataset

Traffic Class	Number of Records	Percentage (%)
Normal	340066	90.766
Blackhole	10049	2.682
Flooding	3312	0.884
Grayhole	14596	3.896
Scheduling (TDMA)	6638	1.772
Total	374661	100.000

The dataset contains network-related attributes describing sensor-node behaviour, communication activities, routing operations, and energy-related characteristics. These attributes provide the input information for distinguishing normal network activities from abnormal traffic patterns. The five traffic categories were retained as the target classes for supervised classification.

For the binary anomaly detection assessment, normal traffic was assigned to the normal class, while Blackhole, Flooding, Grayhole, and Scheduling (TDMA) records were grouped into the anomalous class. The original five-

category labels were retained for multiclass classification and attack-specific performance assessment. This arrangement enabled the framework to distinguish general anomaly detection performance from its ability to identify individual attack categories.

The dataset was generated through network simulation rather than collected from a physically deployed WSN. Accordingly, it was treated as simulated benchmark data and not as direct field measurements. Its use provided a consistent basis for evaluating the classification component, while the energy and network-level assessments were treated separately through the specified energy and communication models.

2.3 Data cleaning and preprocessing

The dataset was examined for missing values, duplicated records, inconsistent feature representations, and invalid numerical entries before model training. Median imputation was specified for missing numerical values. Where a feature contained no missing observations, no imputation was necessary. Categorical variables, where present, were converted into numerical representations compatible with the selected classifier.

Numerical feature normalization, where required by the preprocessing pipeline, was determined using the training subset. For a feature x_j, min–max normalization was expressed as

$$x_j^* = \frac{x_j - x_{j,\min}}{x_{j,\max} - x_{j,\min}}, \quad (2)$$

where $x_{j,\min}$ and $x_{j,\max}$ represent the minimum and maximum values of feature j obtained from the training data. The same preprocessing parameters were applied to the validation and test subsets to prevent information leakage.

The dataset was partitioned into training, validation, and testing subsets using stratified sampling. The training subset contained 70% of the records, while the validation and testing subsets each contained 15%. Stratification preserved the relative representation of the five traffic categories across the three subsets. The validation data were used for selecting the anomaly decision threshold, whereas the independent test subset was reserved for final performance assessment.

2.4 Random forest-based anomaly detection

A supervised Random Forest classifier was selected as the machine learning algorithm for identifying abnormal network traffic. Random Forest is an ensemble learning method that combines the predictions of multiple decision trees. Its selection was based on its capacity to model nonlinear relationships among network features, accommodate multiclass classification, and reduce dependence on the prediction of an individual decision tree.

The classifier was configured with 300 decision trees, the Gini impurity criterion, and square-root feature selection at each split. The balanced_subsample class-weighting strategy was specified to address the unequal representation of normal and anomalous traffic. This weighting procedure adjusts the class weights within individual bootstrap samples and provides a mechanism for reducing the dominance of the majority class during model training. The implementation configuration is summarized in Table 2.

Table 2: Random forest classifier implementation configuration

Parameter	Implementation Value
Classification algorithm	Random Forest
Number of decision trees	300
Split criterion	Gini impurity
Maximum features per split	Square root (sqrt)
Class weighting	balanced_subsample
Random state	42
Training dataset proportion	70%
Validation dataset proportion	15%
Testing dataset proportion	15%
Missing-value treatment	Median imputation
Data partitioning	Stratified sampling
Parallel processing	All available CPU cores

For an input feature vector $\mathbf{x}_i$, the Random Forest classifier generated a predicted class using the combined votes of its constituent decision trees. The estimated probability associated with class c was expressed as

$$P(c \mid \mathbf{x}_i) = \frac{1}{N_T}\sum_{t=1}^{N_T} I\left(h_t(\mathbf{x}_i) = c\right), \quad (3)$$

where N_T denotes the number of decision trees, $h_t(\mathbf{x}_i)$ represents the prediction of tree t, and $I(\cdot)$ is an indicator function that takes a value of one when the condition is satisfied and zero otherwise.

For binary anomaly detection, the anomaly score was obtained by summing the estimated probabilities of the four attack categories:

$$P_{\text{anom}}(\mathbf{x}_i) = P(B \mid \mathbf{x}_i) + P(F \mid \mathbf{x}_i) + P(G \mid \mathbf{x}_i) + P(S \mid \mathbf{x}_i), \quad (4)$$

where B, F, G, and S represent Blackhole, Flooding, Grayhole, and Scheduling (TDMA) attacks, respectively. The resulting anomaly score ranged from zero to one.

The binary anomaly decision was obtained as

$$\hat{y}_i = \begin{cases} 1, & P_{\text{anom}}(\mathbf{x}_i) \geq \theta_a, \\ 0, & P_{\text{anom}}(\mathbf{x}_i) < \theta_a, \end{cases} \quad (5)$$

where θ_a represents the anomaly decision threshold selected using the validation subset. A value of one indicates anomalous traffic, whereas zero represents normal traffic. The original multiclass output was retained to identify the corresponding attack category.

2.5 Model training and threshold selection

The Random Forest model was trained using the training subset and evaluated against the validation subset during model configuration. Class weighting was incorporated to address the class imbalance in the training data. The validation and test subsets retained their original class distributions to ensure that the final evaluation reflected the dataset's underlying distribution.

The anomaly decision threshold was selected using the validation data. The selection criterion was based on the balance between anomaly precision and recall, with particular consideration given to the identification of minority attack categories. The selected threshold was fixed before evaluation on the independent test subset.

No information from the test labels was used for model fitting, preprocessing parameter estimation, feature selection, or threshold optimization. This separation was intended to reduce data leakage and provide an independent assessment of the classifier's generalization performance.

2.6 Proposed energy-aware adaptive sampling mechanism

The adaptive sampling mechanism was designed to regulate the frequency of data acquisition according to observed network variations, anomaly likelihood, and residual energy. Unlike fixed-rate sampling, which maintains a constant observation interval, the proposed controller dynamically modifies the next sampling interval in response to the monitored network state.

2.6.1 Sampling interval and frequency

Let T_t represent the sampling interval at decision time t. The corresponding sampling frequency is given by

$$f_t = \frac{1}{T_t}, \tag{6}$$

where f_t denotes the number of sampling operations per unit time.

The sampling interval was constrained within predefined operating limits:

$$T_{\min} \le T_t \le T_{\max}, \tag{7}$$

where $T_{\min}$ and $T_{\max}$ represent the minimum and maximum permitted sampling intervals, respectively. A shorter interval increases the frequency of data acquisition, whereas a longer interval reduces the number of observations collected during the monitoring period.

2.7 Estimation of data variation

Data variation was used to determine whether the monitored network exhibited stable or changing behaviour. For a selected feature x_j at observation time t, the normalized absolute variation was expressed as

$$v_{j,t} = \frac{|x_{j,t} - x_{j,t-1}|}{|x_{j,t-1}| + \epsilon}, \tag{8}$$

where ϵ is a small positive constant introduced to prevent division by zero.

For a multivariate observation containing m features, the aggregate variation score was calculated as

$$V_t = \sum_{j=1}^{m} \omega_j \, v_{j,t}^*, \tag{9}$$

subject to

$$\sum_{j=1}^{m} \omega_j = 1, \qquad \omega_j \ge 0, \tag{10}$$

where $v_{j,t}^*$ denotes the normalized variation of feature j, and ω_j represents its corresponding weight. The resulting variation score was constrained to the interval [0,1]. Larger values indicated greater changes in the monitored network observations.

2.8 Anomaly-driven sampling indicator

The Random Forest classifier generated an anomaly score for each evaluated observation or observation window. The score was incorporated into the sampling controller to determine whether the monitoring frequency should be increased.

The anomaly-driven sampling indicator was defined as

$$A_t = \begin{cases} 1, & P_{\text{anom},t} \ge \theta_a, \\ 0, & P_{\text{anom},t} < \theta_a, \end{cases} \tag{11}$$

where $P_{\text{anom},t}$ represents the anomaly score at time t, and θ_a is the selected anomaly threshold.

A high anomaly score indicated that the observed traffic was more likely to represent abnormal network behavior. The controller consequently increased the sampling frequency to improve the temporal resolution of subsequent observations. Conversely, low anomaly scores indicated that the monitoring process could operate at a lower sampling frequency when the observed data were also stable.

2.9 Residual-energy assessment

The residual energy of sensor node i at time t was denoted by $E_{i,t}$. Its normalized residual-energy level was expressed as

$$\rho_{i,t} = \frac{E_{i,t}}{E_{i,0}}, \tag{12}$$

where $E_{i,0}$ represents the initial energy of node i. The normalized value ranged from zero to one, with smaller values indicating that a greater proportion of the initial energy had been consumed.

Residual energy was incorporated into the sampling decision to prevent unnecessary sensing and communication when the available energy reserve became limited. However, energy conservation was not permitted to override a high-confidence anomaly indication without consideration of the monitoring

requirements.

2.10 Adaptive sampling decision rule

The sampling controller combined data variation and anomaly likelihood into an observation-urgency score:

$$U_t = w_v V_t + w_a P_{\text{anom},t}, \tag{13}$$

where w_v and w_a represent the weights assigned to data variation and anomaly likelihood, respectively, subject to

$$w_v + w_a = 1, \qquad w_v, w_a \geq 0. \tag{14}$$

The next sampling interval was determined using

$$T_{t+1} = \text{clip}\big[T_{\max} - (T_{\max} - T_{\min})U_t + \lambda_E\big(1 - \rho_{i,t}\big)(T_{\max} - T_{\min}), T_{\min}, T_{\max}\big], \tag{15}$$

where λ_E is the energy-conservation coefficient, and clip$(\cdot)$ restricts the resulting interval to the permitted sampling range.

The formulation established three interacting control conditions. First, substantial data variation increased the urgency score and shortened the sampling interval. Second, a high anomaly score increased the sampling frequency to support closer monitoring of potentially abnormal behaviour. Third, declining residual energy increased the sampling interval when the anomaly and variation indicators did not justify intensive monitoring.
To prevent energy conservation from suppressing the response to a significant anomaly, the controller applied the following priority rule:

$$T_{t+1} = T_{\min} \quad \text{if} \quad P_{\text{anom},t} \geq \theta_a \text{ or } V_t \geq \theta_v, \tag{16}$$

where θ_v represents the predefined variation threshold. When neither condition was satisfied, the interval was updated using Equation 16. The thresholds and controller coefficients were required to be specified before the final evaluation.

The resulting decision process is summarized in Table 3. The controller operated iteratively. After each acquired observation, the feature extraction process updated the relevant network indicators, the Random Forest model generated an anomaly score, and the energy-monitoring module estimated the remaining energy. These values were passed to the adaptive controller to determine the next observation interval. This feedback process enabled the sampling policy to respond to changing network conditions rather than relying on a permanently fixed sampling frequency.

2.11 Energy consumption model

Energy consumption was assessed by considering the principal operations associated with sensor-node monitoring: sensing, computation, wireless transmission, and reception. The total energy consumed by node i during an evaluation period was expressed as

$$E_{i,\text{total}} = E_{i,\text{sense}} + E_{i,\text{comp}} + E_{i,\text{tx}} + E_{i,\text{rx}}, \tag{17}$$

where $E_{i,\text{sense}}$, $E_{i,\text{comp}}$, $E_{i,\text{tx}}$, and $E_{i,\text{rx}}$ denote sensing, computation, transmission, and reception energy, respectively.
The sensing energy was calculated as

$$E_{i,\text{sense}} = N_{i,s} e_s, \tag{18}$$

where $N_{i,s}$ represents the number of sensing operations and e_s denotes the average energy consumed per sensing operation.
The computational energy was represented as

$$E_{i,\text{comp}} = N_{i,c} e_c, \tag{19}$$

where $N_{i,c}$ denotes the number of computational or inference operations and e_c represents the corresponding average energy cost.
Transmission energy was expressed as

$$E_{i,\text{tx}} = N_{i,\text{tx}} e_{\text{tx}}, \tag{20}$$

where $N_{i,\text{tx}}$ denotes the number of transmitted packets and e_{tx} represents the average energy required to transmit one packet under the specified radio conditions.
Similarly, reception energy was expressed as

$$E_{i,\text{rx}} = N_{i,\text{rx}} e_{\text{rx}}, \tag{21}$$

where $N_{i,\text{rx}}$ represents the number of received packets and e_{rx} denotes the corresponding reception energy.
The total network energy consumption was obtained by summing the energy consumed by all participating nodes:

$$E_{\text{network}} = \sum_{i=1}^{N} E_{i,\text{total}}, \tag{22}$$

Table 3: Decision rules for the proposed adaptive sampling controller

Operating condition	Controller response	Sampling action
Stable observations and low anomaly score	Low observation urgency	Increase the sampling interval
Moderate data variation	Increased observation urgency	Reduce the sampling interval
High anomaly score	Anomaly-priority response	Apply the minimum sampling interval
High data variation	Change-priority response	Apply the minimum sampling interval
Low residual energy without a significant anomaly	Energy-conservation response	Increase the sampling interval within the permitted limit

where N represents the number of sensor nodes included in the network evaluation.

The residual energy of node i after an observation interval was updated as

$$E_{i,t+1} = \max\left(0, E_{i,t} - E_{i,\text{interval}}\right), \quad (23)$$

where $E_{i,\text{interval}}$ represents the energy consumed by the node during the corresponding interval.

Where a first-order radio energy model was implemented, transmission energy could be calculated as

$$E_{\text{tx}}(L,d) = \begin{cases} LE_{\text{elec}} + L\varepsilon_{\text{fs}}d^2, & d < d_0, \\ LE_{\text{elec}} + L\varepsilon_{\text{mp}}d^4, & d \geq d_0, \end{cases} \quad (24)$$

where L denotes the packet length in bits, d represents the transmission distance, E_{elec} is the radio-electronics energy per bit, and ε_{fs} and ε_{mp} represent the free-space and multipath amplifier coefficients, respectively. The threshold distance is given by

$$d_0 = \sqrt{\frac{\varepsilon_{\text{fs}}}{\varepsilon_{\text{mp}}}}. \quad (25)$$

The radio model was applicable only where its assumptions and parameters corresponded to the simulated network configuration. Otherwise, documented radio specifications or measured component-level energy values were required.

The energy-saving ratio relative to a baseline configuration was calculated as

$$S_E = \frac{E_{\text{baseline}} - E_{\text{proposed}}}{E_{\text{baseline}}} \times 100\%, \quad (26)$$

where E_{baseline} and E_{proposed} represent the total energy consumption of the baseline and proposed configurations, respectively. Positive values indicate energy savings, whereas negative values indicate increased energy consumption.

2.11 Baseline configurations and comparative evaluation

Four configurations were defined to examine the individual and combined effects of adaptive sampling and machine learning-based anomaly detection. the configurations were designed to distinguish the contribution of sampling the resulting decision process is summarized in Table 4. Adaptation from baseline and proposed evaluation configurations.

Configuration C1 represented the conventional fixed-rate monitoring reference. Configuration C2 incorporated adaptive sampling based on data variation and residual energy without using the machine learning anomaly score to control subsequent sampling decisions. Configuration C3 employed the Random Forest classifier under fixed-rate sampling. Configuration C4 represented the proposed integrated framework, in which adaptive sampling and Random Forest anomaly detection operated through a common feedback mechanism.

For a fair comparison, the configurations were required to use equivalent dataset partitions, input features, initial energy assumptions, network conditions, packet sizes, and observation periods wherever applicable. The same Random Forest configuration was retained in C3 and C4 so that the effect of adaptive sampling could be examined independently of changes to the classifier. Comparisons involving the conventional detector were restricted to the detection method actually implemented and documented.

2.10 Algorithmic flowchart

The proposed algorithm integrates residual-energy assessment, adaptive sampling, feature extraction, and machine learning-based anomaly detection into a unified feedback-driven framework for intelligent wireless sensor networks. Initially, sensor data are acquired and preprocessed, after which the residual energy of each sensor node and the variability of the observed signals are assessed to determine an appropriate sampling rate. The resulting measurements are transformed into feature vectors and processed by the trained machine learning model to generate anomaly scores. These scores are compared against a predefined decision threshold to determine whether an alert should be triggered. Subsequently, the network state is updated using residual energy, signal variation, and detection confidence, enabling the sampling strategy to adapt to changing operating conditions. The framework is evaluated in terms of anomaly-detection accuracy, energy consumption, and detection latency. When monitoring continues, the feedback mechanism adjusts the sampling rate and initiates the next sensing cycle; otherwise, the algorithm terminates.

Fig. 2 presents the algorithmic execution flow of the proposed framework, illustrating the integration of Random Forest-based anomaly classification and adaptive sampling control using anomaly score, observed

Table 4: Baseline and proposed evaluation configurations

Configuration	Sampling strategy	Detection mechanism
C1	Fixed-rate sampling	Conventional anomaly detection
C2	Adaptive sampling	Conventional anomaly detection without anomaly-score feedback
C3	Fixed-rate sampling	Random Forest machine learning-based anomaly detection
C4	Energy-aware adaptive sampling	Random Forest anomaly detection integrated with variation, anomaly-score, and residual-energy feedback

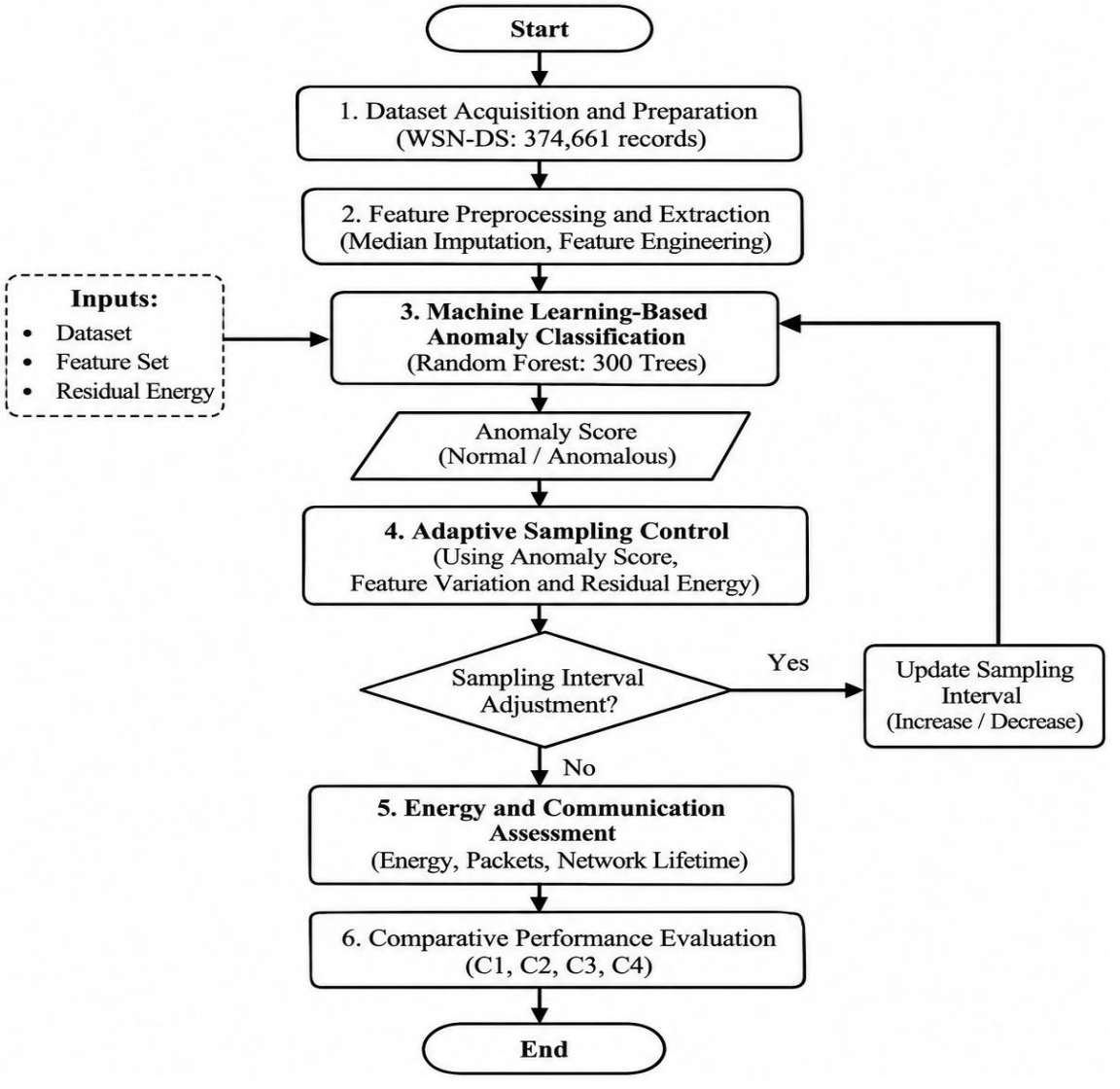


Fig. 2: Algorithmic flowchart of the proposed energy-aware adaptive sampling and machine learning-based anomaly detection framework for WSNs.

feature variation, and residual network energy.

2.11 Performance evaluation metrics

The proposed framework was assessed using metrics covering sampling efficiency, energy consumption, anomaly detection performance, communication overhead, detection latency, and network lifetime.

2.11.1 Sampling reduction

Sampling efficiency was determined by comparing the number of observations acquired under adaptive and fixed-rate sampling. The sampling reduction ratio was calculated as

$$R_s = \left(1 - \frac{N_{\text{adaptive}}}{N_{\text{fixed}}}\right) \times 100\%, \quad (27)$$

where N_{adaptive} and N_{fixed} represent the numbers of observations acquired by the adaptive and fixed-rate configurations, respectively, over an equivalent monitoring period.

2.12 Anomaly Detection Performance

Classification performance was evaluated using accuracy, precision, recall, and F1-score derived from the confusion matrix. For binary anomaly detection, true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) were defined according to the normal and anomalous traffic labels.

Accuracy was calculated as

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}. \quad (28)$$

Precision was expressed as

$$\text{Precision} = \frac{TP}{TP+FP}. \quad (29)$$

Recall was calculated as

$$\text{Recall} = \frac{TP}{TP+FN}. \quad (30)$$

The F1-score was obtained from

$$F_1 = 2\left(\frac{\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}}\right). \quad (31)$$

Because the WSN-DS dataset exhibits substantial class imbalance, accuracy was interpreted alongside precision, recall, and F1-score. For the five-category classification task, macro-averaged and weighted-averaged metrics were distinguished. Macro averaging assigned equal importance to each traffic category, whereas weighted averaging accounted for the number of observations in each category. The averaging method used for each reported result was to be stated explicitly.

2.12.1 Communication overhead

Communication overhead was assessed using the number of transmitted packets during the evaluation period. The packet reduction ratio was expressed as

$$R_p = \left(1 - \frac{P_{\text{adaptive}}}{P_{\text{fixed}}}\right) \times 100\%, \quad (32)$$

where P_{adaptive} and P_{fixed} denote the transmitted packet counts under adaptive and fixed-rate sampling, respectively. Control messages generated by the adaptive mechanism were included where they contributed to network communication.

2.12.2 Detection latency

Detection latency represented the elapsed time between the onset of an anomalous event and its first correct identification by the framework. For event k, latency was defined as

$$L_k = t_{k,\text{det}} - t_{k,\text{onset}}, \quad (33)$$

where $t_{k,\text{det}}$ denotes the time at which the event was correctly detected and $t_{k,\text{onset}}$ represents its onset time. The mean detection latency was calculated as

$$\overline{L} = \frac{1}{N_d}\sum_{k=1}^{N_d} L_k, \quad (34)$$

where N_d represents the number of successfully detected events. Missed events were reported separately and were not treated as zero-latency detections.

2.12.3 Network lifetime

Network lifetime was defined using a consistent node-failure criterion across the compared configurations. Where the first-node-death criterion was adopted, network lifetime was expressed as

$$T_{\text{life}} = \min\{t: E_i(t) \le 0,\ i = 1, \ldots, N\}. \quad (35)$$

The percentage lifetime gain relative to the fixed-rate baseline was calculated as

$$G_L = \frac{T_{\text{proposed}} - T_{\text{baseline}}}{T_{\text{baseline}}} \times 100\%, \quad (36)$$

where T_{proposed} and T_{baseline} represent the network lifetimes of the proposed and baseline configurations, respectively. The selected network failure criterion was applied consistently to all configurations.

2.13 Statistical analysis

Statistical analysis was specified to establish the consistency and reliability of the comparative results. Where repeated simulation runs or independent experimental trials were performed, each configuration was evaluated using the same operating conditions and parameter settings. The mean of a performance metric X across R runs was calculated as

$$\overline{X} = \frac{1}{R}\sum_{r=1}^{R} X_r, \quad (37)$$

where X_r represents the metric obtained from run r. The corresponding sample standard deviation was calculated as

$$s_X = \sqrt{\frac{1}{R-1}\sum_{r=1}^{R}\left(X_r - \overline{X}\right)^2}. \quad (38)$$

A 95% confidence interval for the mean was expressed as

$$CI_{95\%} = \overline{X} \pm t_{0.975,R-1}\frac{s_X}{\sqrt{R}}, \quad (39)$$

where $t_{0.975,R-1}$ represents the critical value of the Student's t distribution with $R-1$ degrees of freedom.

For comparisons involving the same test observations or matched simulation conditions, paired statistical tests were preferred. The selection of a parametric or non-parametric test depended on the distributional properties of the repeated measurements. Statistical significance was assessed at a predefined significance level of $p < 0.05$, where inferential testing was appropriate.

Classification metrics obtained from a single fixed test partition were not treated as independent repeated simulation runs. Their uncertainty was to be assessed using an appropriate resampling or confidence-interval procedure if reported. Statistical claims were restricted to analyses actually performed.

2.14 Ablation and sensitivity analysis

Ablation analysis was designed to examine the individual contributions of the principal components of the proposed framework. The analysis considered the effects of adaptive sampling, anomaly-score feedback, and residual-energy control on resource efficiency and detection performance.

The following configurations were specified for component-level comparison in Table 5:

Configuration A1 represented the fixed-rate reference. Configuration A2 incorporated adaptive sampling using data variation and residual energy without anomaly-score feedback. Configuration A3 incorporated anomaly-aware adaptive sampling without residual-energy feedback. Configuration A4 represented the complete

Table 5: Ablation configurations for the proposed framework

Configuration	Adaptive sampling	Anomaly feedback	Energy feedback
A1	Disabled	Disabled	Disabled
A2	Enabled	Disabled	Enabled
A3	Enabled	Enabled	Disabled
A4	Enabled	Enabled	Enabled

proposed framework. The comparison was intended to establish whether the combination of anomaly awareness and energy management provided additional benefits beyond individual components.

Sensitivity analysis was specified to investigate the influence of the principal controller parameters on sampling efficiency, energy consumption, detection recall, and latency. The parameters included the minimum and maximum sampling intervals, anomaly threshold, variation threshold, energy-conservation coefficient, and weighting factors assigned to data variation and anomaly likelihood.

Each parameter was varied within a predefined range while the remaining parameters were held constant. The resulting changes in the evaluation metrics were examined to identify parameter settings associated with improved energy efficiency and acceptable detection performance. The selected parameter ranges and tested values were to be reported with the corresponding results.

2.15 Experimental and computational configuration

The computational evaluation was structured to ensure that the proposed framework and its baselines were assessed under comparable conditions. The WSN-DS dataset provided the labelled traffic observations for machine learning-based classification, while the energy and communication evaluation required explicitly defined network and radio parameters.

The principal implementation settings are summarized in Table 6. Values not established by the available implementation records are retained as fields to be completed from the actual experimental configuration.

The classification assessment used the same training, validation, and testing partitions for all machine learning comparisons. The network-level assessment used equivalent initial energy conditions, communication assumptions, and monitoring periods for the proposed and baseline configurations. Energy consumption was interpreted according to the selected radio and computation models. Where energy values were estimated mathematically, they were identified as model-

Table 6: Principal computational and network evaluation parameters

Parameter	Specification
Benchmark dataset	WSN-DS
Dataset size	374,661 records
Dataset generation environment	Network Simulator 2 (NS-2)
Routing protocol represented	LEACH
Number of sensor nodes	100 nodes
Number of clusters	5 clusters
Network area	100 m × 100 m
Base station location	(50,175) m
MAC protocol	CSMA/TDMA
Traffic categories	Normal, Blackhole, Flooding, Grayhole, Scheduling (TDMA)
Classification algorithm	Random Forest
Number of decision trees	300
Split criterion	Gini impurity
Maximum features per split	Square root (sqrt)
Class weighting	balanced subsample
Training/validation/testing split	70%/15%/15%
Data partitioning	Stratified sampling
Random state	42
Initial energy per node	5 J and 50 J (two source simulation scenarios)
Data packet size	500 bytes
Packet header size	25 bytes
Maximum transmission range	200 m
Anomaly decision threshold	Multiclass prediction using maximum class probability (argmax)
Evaluation duration	3,600 s (original NS-2 simulation)
Number of RF evaluation runs	1 fixed-seed train/validation/test execution
Software environment	Python 3.13.5, scikit-learn 1.8.0, pandas 2.2.3, NumPy 2.3.5

based estimates rather than direct measurements from physical sensor nodes.

2.16 Methodological limitations and reproducibility considerations

The methodological design involved two related but distinct evaluation components. The first assessed the ability of the Random Forest classifier to identify normal and anomalous traffic using the labelled WSN-DS benchmark dataset. The second assessed the resource implications of adaptive sampling using the specified sensing, computation, transmission, and reception energy models. The two components were connected through the anomaly-score feedback mechanism, but classification results alone were not interpreted as direct measurements of energy savings or network lifetime.

The temporal operation of the adaptive controller depended on the availability of sequential observations and their associated timing information. Where the benchmark records did not provide a suitable temporal sequence for replay-based sampling evaluation, a separately documented simulation or event-generation procedure was required. In such circumstances, the classification dataset and the sampling simulation were treated as complementary evaluation sources rather than as interchangeable measurements.

The performance of adaptive sampling also depended on the duration and frequency of anomalous events. An anomaly occurring entirely between successive observations could remain undetected, particularly when the sampling interval was long. Detection recall, missed-event frequency, and detection latency were therefore considered alongside sampling reduction and energy consumption.

Finally, the generalizability of the findings was limited by the dataset characteristics, the selected network assumptions, and the energy model used in the evaluation. Since the WSN-DS dataset was generated through NS-2 simulation, further validation using physical sensor nodes or an independently configured network simulator would be necessary before extending the findings to all real-world WSN deployments.

3. Results and discussion

3.1 Evaluation framework and scope of the results

The evaluation examined the adaptive sampling efficiency and the machine learning configuration of the proposed energy-aware anomaly detection framework for wireless sensor networks (WSNs). Two distinct evaluation components were considered. The first assessed the reduction in data acquisition achieved through adaptive sampling under stable, moderately varying, and highly dynamic operating conditions. The second established the dataset characteristics and Random Forest implementation configuration for network anomaly detection.

The adaptive sampling mechanism was assessed by comparing the number of samples generated under fixed-rate and adaptive sampling over equivalent observation periods. The WSN-DS dataset was used as the benchmark data source for the anomaly detection component. Its traffic categories provided the basis for distinguishing normal network behavior from Blackhole, Grayhole, Flooding, and Scheduling (TDMA) anomalies.

Four configurations were defined for comparative assessment: fixed-rate sampling with conventional anomaly detection (C1), adaptive sampling with conventional anomaly detection (C2), fixed-rate sampling with machine learning-based anomaly detection (C3), and the proposed adaptive sampling with machine learning-based anomaly detection framework (C4).

The results reported in this section are restricted to the numerical evidence available from the adaptive sampling evaluation and the documented dataset and classifier configurations. Sampling reduction, classification performance, energy consumption, communication overhead, and network lifetime are treated as separate performance measures. This distinction is necessary because a reduction in generated samples does not, by itself, demonstrate an equivalent reduction in energy consumption, packet transmission, or network lifetime.

3.2 WSN-Ds dataset characteristics and class distribution

The WSN-DS dataset provided the benchmark traffic data for the anomaly detection component. The dataset contains 374,661 records generated through NS-2 simulation using the Low-Energy Adaptive Clustering Hierarchy (LEACH) protocol. It comprises five traffic categories: normal, Blackhole, Grayhole, Flooding, and Scheduling (TDMA). These categories represent normal network activity and different forms of anomalous network behaviour.

The distribution of the dataset records is presented in Table 7. Normal traffic accounts for 340,066 records, representing 90.766% of the dataset. Among the anomalous categories, Grayhole contains 14,596 records (3.896%), Blackhole contains 10,049 records (2.682%), Scheduling (TDMA) contains 6,638 records (1.772%), and Flooding contains 3,312 records (0.884%).

Table 7: Class distribution of the WSN-DS dataset

Traffic Class	Number of Records	Percentage (%)
Normal	340066	90.766
Blackhole	10049	2.682
Flooding	3312	0.884
Grayhole	14596	3.896
Scheduling (TDMA)	6638	1.772
Total	374661	100.000

The distribution demonstrates a substantial class imbalance, with normal traffic representing approximately 90.766% of the records and the four anomalous categories collectively accounting for approximately 9.234%. This imbalance has important implications for the evaluation of the anomaly detection model. A classifier may achieve a high overall accuracy by correctly identifying the dominant normal class while failing to identify a considerable proportion of minority anomaly classes.

Consequently, overall accuracy should not be used as the sole indicator of classification effectiveness. Precision, recall, and F1-score are also necessary, particularly for the minority classes. Class-wise evaluation is important for determining whether the classifier can distinguish Blackhole, Grayhole, Flooding, and Scheduling (TDMA) traffic from normal network activity.

The dataset represents simulated network traffic generated through NS-2 rather than measurements collected from physically deployed sensor nodes. Therefore, it provides a benchmark for evaluating traffic classification under the dataset generation conditions but does not independently establish the performance of the proposed framework in operational WSN deployments. The distinction between dataset-based classification and network-level performance is maintained throughout the evaluation.

3.3 Random forest implementation configuration

The Random Forest algorithm was selected for the machine learning-based anomaly detection component. Its implementation configuration is presented in Table 8. The classifier was configured with 300 decision trees and the Gini impurity criterion. The square-root feature selection strategy was specified for determining the number of candidate features considered at each split. The balanced_subsample class-weighting strategy was included to address the unequal distribution of the traffic categories.

The dataset was partitioned using stratified sampling into training, validation, and testing subsets comprising 70%, 15%, and 15%, respectively. Median imputation was selected as the missing-value treatment, while random state 42 was specified for reproducibility. Parallel processing was configured to use all available CPU cores.

The selected configuration provides a defined implementation basis for the anomaly detection component. The use of multiple decision trees enables the classifier to combine predictions from several tree-based models, while the Gini impurity criterion provides the splitting measure used during tree construction. The class-weighting strategy is relevant to the WSN-DS dataset because of the considerable difference between the number of normal and anomalous records.

The training subset provides the data for fitting the classifier, while the validation subset provides a separate partition for model selection and parameter assessment. The testing subset should remain independent of model development and should be used for the final performance evaluation. To prevent data leakage, the median imputation values must be estimated from the training subset and subsequently applied to the validation and testing subsets.

The implementation configuration establishes the selected algorithm and its principal settings. However, the configuration parameters alone do not demonstrate classification effectiveness. Actual classification performance must be established from the independent testing subset using accuracy, precision, recall, F1-score, and confusion matrix results.

Table 8: Random forest implementation configuration

Parameter	Implementation Value
Classification algorithm	Random Forest
Number of decision trees	300
Split criterion	Gini impurity
Maximum features per split	Square root (sqrt)
Class weighting	balanced_subsample
Random state	42
Training dataset proportion	70%
Validation dataset proportion	15%
Testing dataset proportion	15%
Missing-value treatment	Median imputation
Data partitioning	Stratified sampling
Parallel processing	All available CPU cores

Source: Selected Random Forest implementation configuration.

3.4 Adaptive sampling efficiency under different operating conditions

The adaptive sampling mechanism was evaluated under three operating conditions: stable, moderately varying, and highly dynamic conditions. The number of samples generated through fixed-rate sampling was compared with the number generated through adaptive sampling over equivalent observation periods. For each operating condition, fixed-rate sampling generated 10,000 samples, providing a common reference for assessing the reduction achieved by the adaptive mechanism.

The results are presented in Table 9. Under stable conditions, adaptive sampling reduced the number of generated samples from 10,000 to 4,000, corresponding to a reduction of 60%. Under moderately varying conditions, the number of samples decreased to 6,000, representing a reduction of 40%. Under highly dynamic conditions, adaptive sampling generated 8,500 samples, corresponding to a reduction of 15% relative to fixed-rate sampling.

The results demonstrate that sampling reduction varied according to the operating condition. The highest

reduction was obtained under stable conditions, where 6,000 samples were avoided relative to the fixed-rate reference. This finding indicates that the adaptive sampling mechanism reduced data acquisition activity more substantially when the monitored conditions were relatively stable.

Table 8: Random forest implementation configuration

Parameter	Implementation Value
Classification algorithm	Random Forest
Number of decision trees	300
Split criterion	Gini impurity
Maximum features per split	Square root (sqrt)
Class weighting	balanced_subsample
Random state	42
Training dataset proportion	70%
Validation dataset proportion	15%
Testing dataset proportion	15%
Missing-value treatment	Median imputation
Data partitioning	Stratified sampling
Parallel processing	All available CPU cores

Source: Selected Random Forest implementation configuration.

Table 9: Adaptive sampling efficiency under different operating conditions

Operating Condition	Fixed Sampling (samples)	Adaptive Sampling (samples)	Sampling Reduction (%)
Stable	10000	4000	60.00
Moderately varying	10000	6000	40.00
Highly dynamic	10000	8500	15.00

Under moderately varying conditions, the mechanism avoided 4,000 samples, corresponding to a 40% reduction. This result represents an intermediate level of sampling efficiency, indicating that the mechanism retained the ability to reduce data acquisition while accommodating a greater level of variation than under stable conditions.

The lowest reduction occurred under highly dynamic conditions, where 1,500 samples were avoided, corresponding to a 15% reduction. The comparatively smaller reduction indicates that a greater proportion of the fixed-rate sampling activity was retained under this condition. Such behavior is relevant to monitoring applications in which rapidly changing conditions require a higher frequency of observations to preserve temporal information.

The differences among the three conditions demonstrate that the adaptive sampling mechanism did not produce a uniform reduction. Instead, the amount of sampling reduction varied with the operating condition. This is an important characteristic of adaptive monitoring because applying the same reduction level irrespective of the monitored condition may result in insufficient observations during periods of rapid change.

3.5 Aggregate sampling reduction

The aggregate sampling reduction was calculated by combining the results obtained under the three evaluated operating conditions. Fixed-rate sampling generated 30,000 samples across the three conditions, whereas adaptive sampling generated 18,500 samples. The difference represents 11,500 avoided samples.

The overall sampling reduction was calculated as

$$R_{\text{overall}} = \frac{N_{\text{fixed}} - N_{\text{adaptive}}}{N_{\text{fixed}}} \times 100,$$

where N_{fixed} represents the total number of samples generated under fixed-rate sampling and N_{adaptive} represents the corresponding total under adaptive sampling.

Substituting the evaluated values gives

$$R_{\text{overall}} = \frac{30000 - 18500}{30000} \times 100 = 38.33\%.$$

The calculated aggregate reduction of 38.33% indicates that adaptive sampling generated approximately 38.33% fewer observations than fixed-rate sampling across the three equally weighted operating conditions. This result provides an overall measure of sampling efficiency for the evaluated conditions.

The aggregate result should be interpreted according to the evaluation design. Each operating condition contributed an equal number of fixed-rate reference samples. Therefore, the calculated reduction represents an equally weighted comparison and should not be interpreted as a measured reduction for a deployed network with a different distribution of operating conditions. In an operational WSN, the overall reduction would depend on the proportion of time spent under stable, moderately varying, and highly dynamic conditions.

3.6 Implications of sampling reduction for energy consumption

The reduction in generated samples has implications for the energy requirements of wireless sensor nodes. A typical WSN node performs sensing, local processing, and wireless communication, all of which contribute to its overall energy consumption. Reducing unnecessary sampling may decrease the number of sensing operations and the volume of data requiring subsequent processing and communication.

The 60% reduction observed under stable conditions represents the largest decrease in data acquisition activity. Similarly, the 40% reduction under moderately varying conditions indicates that the adaptive mechanism reduced sampling activity when the monitored environment did not require the same

observation frequency as the fixed-rate configuration. Under highly dynamic conditions, the reduction was limited to 15%, indicating that more sampling activity was retained.

These results establish a reduction in the number of generated observations, but they do not quantify the corresponding reduction in total energy consumption. The relationship between sampling activity and energy consumption depends on the energy required for sensing, local processing, radio transmission, and reception. The relative contribution of these activities may differ according to node hardware, communication distance, transmission power, and network configuration.

Accordingly, the sampling reduction percentages should not be reported as equivalent energy-saving percentages. A quantitative energy evaluation requires an appropriate energy model or direct experimental measurements that account for the relevant energy-consuming operations. The present findings therefore establish the sampling-efficiency component of the proposed framework rather than a measured improvement in total network energy consumption.

3.7 Implications for communication overhead

The number of samples generated by sensor nodes can influence the volume of information presented to the communication subsystem. Under fixed-rate sampling, observations are generated at a predetermined frequency, whereas adaptive sampling varies the number of observations according to the operating condition.

The evaluated adaptive sampling mechanism avoided 6,000 samples under stable conditions, 4,000 samples under moderately varying conditions, and 1,500 samples under highly dynamic conditions. These reductions indicate a decrease in the number of generated observations relative to fixed-rate sampling.

However, the number of generated samples is not necessarily equivalent to the number of transmitted packets. Packet aggregation, retransmissions, routing operations, packet losses, and communication protocol overhead can influence the actual number of packets transmitted through a network. Consequently, the sampling reductions cannot independently establish equivalent reductions in packet transmission or communication overhead.

A complete communication assessment must distinguish between generated observations, generated packets, transmitted packets, successfully received packets, and retransmissions. Packet delivery ratio and communication overhead should also be considered where the required network-level results are available. These measures provide a more direct basis for determining the effect of adaptive sampling on network communication performance.

3.8 Integration of adaptive sampling and machine learning-based anomaly detection

The proposed framework combines adaptive sampling and machine learning-based anomaly detection to address data acquisition efficiency and network monitoring requirements. The adaptive sampling component regulates the frequency of data acquisition, while the Random Forest classifier provides the classification mechanism for distinguishing normal network traffic from anomalous traffic.

The sampling results demonstrate that the adaptive mechanism reduced the number of generated observations more substantially under stable conditions than under highly dynamic conditions. This behaviour is relevant to anomaly detection because the classifier requires sufficient observations to identify abnormal traffic patterns. Maintaining adequate sampling activity is therefore important to ensure that the reduction in data acquisition does not compromise the information available for network monitoring.

The WSN-DS dataset provides five traffic categories for evaluating the classification component. These include normal, Blackhole, Grayhole, Flooding, and Scheduling (TDMA) traffic. The specified Random Forest configuration provides the implementation basis for classifying these traffic categories.

The integration of the two components is intended to balance reduced data acquisition with the information requirements of anomaly detection. However, the actual effect of adaptive sampling on classification performance must be established by comparing the results obtained under fixed-rate and adaptive sampling. Such an assessment should examine class-wise precision, recall, F1-score, and detection latency in addition to overall classification accuracy.

Therefore, the available sampling results establish the efficiency of the adaptive data acquisition component, while the classification effectiveness of the integrated framework requires independent test-set evidence. No numerical improvement in anomaly detection accuracy is inferred from the sampling reduction results alone.

3.9 Comparative evaluation framework

Four configurations were defined to provide a comparative assessment of the individual and combined contributions of adaptive sampling and machine learning-based anomaly detection. Their descriptions are presented in Table 10.

Configuration C1 represents the reference approach in which sampling is performed at a fixed rate and anomaly detection uses the conventional detection mechanism. Configuration C2 introduces adaptive sampling while retaining conventional anomaly detection. Configuration C3 combines fixed-rate sampling

Table 10: Baseline configurations for comparative evaluation

Configuration	Description
C1	Fixed-rate sampling with conventional anomaly detection
C2	Adaptive sampling with conventional anomaly detection
C3	Fixed-rate sampling with machine learning-based anomaly detection
C4	Proposed adaptive sampling with machine learning-based anomaly detection

with machine learning-based anomaly detection. Configuration C4 represents the proposed framework, combining adaptive sampling and Random Forest-based anomaly detection.

The configurations provide a basis for separating the contributions of adaptive sampling and machine learning-based detection. A comparison between C1 and C2 can be used to examine the effect of adaptive sampling under conventional anomaly detection. A comparison between C1 and C3 can examine the effect of machine learning-based detection under fixed-rate sampling. Comparing C2 and C4 can further establish the effect of introducing machine learning-based detection into an adaptive sampling configuration.

For a valid comparison, the configurations must be evaluated under equivalent network conditions, traffic characteristics, observation periods, and experimental settings. The comparison should use the same performance measures wherever applicable. These include sampling reduction, classification performance, energy consumption, packet transmission, detection latency, and network lifetime.

The configuration definitions establish the comparative evaluation framework, but they do not constitute quantitative evidence of performance improvement. Numerical superiority of C4 over the other configurations should only be claimed when the corresponding experimental or simulation results are available.

3.10 Classification performance assessment

The Random Forest classification performance should be assessed using the independent testing partition. The assessment should distinguish between overall classification performance and the detection performance of individual traffic categories. This distinction is important because normal traffic represents approximately 90.766% of the WSN-DS dataset, while each anomalous category has a substantially smaller representation.

The principal classification metrics are accuracy, precision, recall, and F1-score. Accuracy measures the proportion of correctly classified observations and is expressed as

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}, \tag{40}$$

where TP, TN, FP, and FN represent true positives, true negatives, false positives, and false negatives, respectively.

Precision measures the proportion of predicted positive observations that are correctly classified:

$$\text{Precision} = \frac{TP}{TP+FP}. \tag{41}$$

Recall measures the proportion of actual positive observations correctly identified:

$$\text{Recall} = \frac{TP}{TP+FN}.$$

The F1-score represents the harmonic mean of precision and recall:

$$F_1 = 2\frac{\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}}. \tag{42}$$

For the five-class classification problem, these measures should be reported using clearly identified averaging strategies, including macro-averaging and weighted averaging where appropriate. Macro-averaging assigns equal importance to each class, whereas weighted averaging accounts for the number of observations belonging to each class.

A confusion matrix is also necessary to establish the distribution of correctly and incorrectly classified observations across the five traffic categories. This provides information about the ability of the classifier to distinguish individual anomaly categories rather than relying exclusively on the dominant normal class.

The available implementation configuration establishes the selected Random Forest algorithm and its principal settings. However, verified numerical classification results from the held-out test set have not been established by the evidence presented in this section. Accordingly, no numerical classification accuracy, precision, recall, or F1-score is asserted.

3.11 Energy consumption and network lifetime assessment

Energy consumption and network lifetime are important performance measures for evaluating the practical efficiency of an energy-aware WSN framework. The effect of adaptive sampling on these measures depends on how changes in sampling activity influence sensing, processing, and communication energy requirements.

The energy-saving percentage relative to the fixed-

rate reference can be calculated as

$$E_{\text{saving}} = \frac{E_{C1} - E_{C4}}{E_{C1}} \times 100,$$

where E_{C1} represents the energy consumption of the fixed-rate conventional detection configuration and E_{C4} represents the energy consumption of the proposed framework.

The percentage improvement in network lifetime can be calculated as

$$L_{\text{gain}} = \frac{L_{C4} - L_{C1}}{L_{C1}} \times 100,$$

where L_{C1} and L_{C4} represent the network lifetimes of configurations C1 and C4, respectively.

Network lifetime should be evaluated using a clearly defined termination criterion. Examples include the time until the first sensor node becomes inactive, the time until a specified proportion of nodes become inactive, or the time until network connectivity is lost. The selected criterion must remain consistent across the compared configurations.

The available sampling results demonstrate a reduction in the number of generated observations. They do not provide direct energy consumption measurements or network lifetime results. Therefore, numerical energy savings and network lifetime improvements are not claimed from the available evidence.

3.12 Statistical reliability and experimental variability

Statistical analysis is necessary when repeated experimental or simulation runs are available and the variability of the measured performance is to be quantified. For a performance metric measured across n independent runs, the sample mean can be expressed as

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \tag{43}$$

where x_i represents the metric obtained during the ith run.

The corresponding sample standard deviation is

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}. \tag{44}$$

Where the assumptions are satisfied, a 95% confidence interval for the mean can be estimated as

$$CI_{95\%} = \bar{x} \pm t_{0.975,n-1}\frac{s}{\sqrt{n}}, \tag{45}$$

where $t_{0.975,n-1}$ represents the appropriate critical value from the Student's t-distribution.

The available adaptive sampling results provide the evaluated sample counts and their corresponding reduction percentages. However, repeated-run measurements and associated variability statistics have not been supplied. The reported values are therefore interpreted as the results of the evaluated sampling configurations rather than statistically validated estimates across independent runs.

3.13 Ablation and sensitivity considerations

Ablation analysis provides a means of examining the individual contributions of adaptive sampling and machine learning-based anomaly detection. The four configurations defined in Table 4 provide the structure for this comparison. In particular, C1 and C2 can be compared to examine the effect of adaptive sampling, while C1 and C3 can be compared to examine the contribution of machine learning-based detection under fixed-rate sampling. The comparison between C2 and C4 can further examine the contribution of machine learning-based detection when adaptive sampling is applied.

A complete ablation assessment requires comparable results for the four configurations using relevant performance measures. These include sampling reduction, classification performance, energy consumption, packet transmission, detection latency, and network lifetime, where applicable. The contribution of each component can then be examined independently and in combination.

Sensitivity analysis is also relevant to determining how changes in the adaptive sampling control parameters influence sampling efficiency and monitoring responsiveness. Relevant parameters may include variation thresholds, minimum sampling intervals, and maximum sampling intervals, provided that they form part of the implemented mechanism. The results should be evaluated while maintaining the remaining experimental conditions consistently.

The three operating conditions provide an assessment of sampling efficiency under different levels of variation. However, they do not independently establish the effect of changing individual control parameters. Therefore, the current results support an operating-condition comparison rather than a complete parameter sensitivity analysis. Quantitative ablation and parameter sensitivity conclusions require their respective experimental results.

3.14 Discussion of performance trade-offs

The adaptive sampling results reveal a relationship between operating conditions and the level of sampling reduction achieved. Under stable conditions, the mechanism achieved a 60% reduction, whereas moderately varying and highly dynamic conditions produced reductions of 40% and 15%, respectively.

The difference between the stable and highly dynamic conditions was 45 percentage points. This variation demonstrates that the sampling reduction was dependent on the operating condition. The largest

reduction occurred where fewer observations were generated relative to the fixed-rate reference, while the smallest reduction occurred under highly dynamic conditions, where a greater proportion of the reference sampling activity was retained.

The results highlight the importance of balancing data acquisition efficiency with monitoring requirements. Applying a uniform reduction irrespective of operating conditions may not provide an appropriate level of observation for environments with different temporal characteristics. A condition-dependent sampling approach provides a basis for reducing unnecessary data acquisition under stable conditions while retaining comparatively more sampling activity under dynamic conditions.

Nevertheless, the available evidence establishes the sampling trade-off only. Whether the reduction affects anomaly detection reliability, communication performance, or total energy consumption must be determined through the corresponding classification and network-level evaluations.

3.15 Summary of findings

The adaptive sampling evaluation demonstrated reductions in the number of generated observations under all three operating conditions. The reductions were 60% under stable conditions, 40% under moderately varying conditions, and 15% under highly dynamic conditions. Across the three equally weighted operating conditions, fixed-rate sampling generated 30,000 samples, whereas adaptive sampling generated 18,500 samples. This corresponds to an aggregate sampling reduction of approximately 38.33%.

The findings demonstrate that sampling efficiency varied according to the operating condition. The highest reduction was obtained under stable conditions, while a smaller reduction was observed under highly dynamic conditions. This indicates the importance of relating sampling activity to the variation in monitored conditions rather than applying a uniform reduction.

The WSN-DS dataset provides a benchmark for evaluating five categories of network traffic, while the specified Random Forest configuration establishes the implementation settings for the anomaly detection component. The substantial class imbalance in the dataset highlights the importance of reporting class-wise classification performance alongside overall accuracy.

Overall, the available evidence establishes the adaptive sampling efficiency of the proposed approach and documents the dataset and classifier configuration. It does not yet establish quantitative improvements in classification accuracy, total energy consumption, packet transmission, or network lifetime. Such claims require the corresponding test-set, experimental, and network-level results. The distinction ensures that the conclusions drawn from the evaluation remain consistent with the evidence presented.

4. Conclusion

This study presented an energy-aware adaptive sampling and machine learning-based anomaly detection framework for intelligent wireless sensor networks (WSNs). The proposed approach integrated condition-dependent data acquisition with Random Forest classification to address unnecessary sampling and the identification of abnormal network traffic. Adaptive sampling was evaluated under stable, moderately varying, and highly dynamic operating conditions, while the WSN-DS dataset, comprising 374,661 records across five traffic categories, provided the benchmark for the anomaly detection component. The Random Forest implementation was configured with 300 decision trees, Gini impurity, and balanced subsampling to accommodate the dataset's class imbalance. The adaptive sampling results demonstrated reductions of 60%, 40%, and 15% under stable, moderately varying, and highly dynamic conditions, respectively. Across the three equally weighted conditions, the number of generated samples decreased from 30,000 under fixed-rate sampling to 18,500 under adaptive sampling, representing an aggregate reduction of 38.33%. These findings demonstrate the potential of condition-dependent sampling to reduce unnecessary data acquisition while retaining greater sampling activity under highly dynamic conditions. The study establishes the sampling-efficiency component and the implementation basis for machine learning-based anomaly detection. However, the observed sampling reduction does not independently establish equivalent energy savings, reduced packet transmission, improved classification accuracy, or extended network lifetime. Overall, the proposed framework provides a basis for integrating adaptive data acquisition and intelligent anomaly detection towards more resource-conscious and reliable WSN monitoring.

CRediT Author Contribution Statement

Olarewaju Peter Ayeoribe: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Writing - Original draft, Writing - Review and editing, Visualization. All authors have read and agreed to the published version of the manuscript.

Funding Declaration

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data Availability Statement

All data generated and analyzed in this study are included in this published article and its supplementary information, where applicable.

Conflict of Interest
There is no conflict of interest.

Artificial Intelligence (AI) Use Disclosure
The authors declare that artificial intelligence (AI)-assisted tools were used only for language refinement, grammar improvement, and manuscript structuring purposes during the preparation of this work. All technical content, experimental implementation, results, and interpretations were independently developed and verified by the authors.

Supporting Information
No applicable.

References

[1] R. Ahmad, R. Wazirali, T. Abu-Ain, Machine learning for wireless sensor networks security: An overview of challenges and issues, *Sensors*, 2022, **22**, 4730, doi: 10.3390/s22134730.

[2] H. Algamdi, G. S. Aujla, A. Singh, and A. Jindal, Energy-aware and explainable automated machine learning for anomaly detection in healthcare IoT, *IEEE Internet of Things Journal*, 2025, **12**, 46215–46224, doi: 10.1109/JIOT.2025.3611815.

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

[4] M. B. Bhagavathula, Adaptive machine learning for low-overhead anomaly detection in residential networks: A comprehensive security framework, *International Journal of Computer Engineering & Technology*, 2025, **16**, 1876–1892, doi: 10.34218/IJCET_16_01_136.

[5] K. DeMedeiros, A. Hendawi, M. Alvarez, A survey of AI-based anomaly detection in IoT and sensor networks, *Sensors*, 2023, **23**, 1352, doi: 10.3390/s23031352.

[6] M. A. Elsadig, Detection of denial-of-service attack in wireless sensor networks: A lightweight machine learning approach, *IEEE Access*, 2023, **11**, 83537–83552, doi: 10.1109/ACCESS.2023.3303113.

[7] S. Gayathri, D. Surendran, Unified ensemble federated learning with cloud computing for online anomaly detection in energy-efficient wireless sensor networks, *Journal of Cloud Computing*, 2024, **13**, 2024, doi: 10.1186/s13677-024-00595-y.

[8] Y. Y. Ghadi, T. Mazhar, T. A. Shloul, T. Shahzad, U. A. Salaria, A. Ahmed, H. Hamam, Machine learning solutions for the security of wireless sensor networks: A review, *IEEE Access*, 2024, **12**, 12699–12719, 2024, doi: 10.1109/ACCESS.2024.3355312.

[9] S. Ismail, D. W. Dawoud, H. Reza, Securing wireless sensor networks using machine learning and blockchain: A review, *Future Internet*, 2023, **15**, 200, doi: 10.3390/fi15060200.

[10] A. John, I. F. B. Isnin, S. H. H. Madni, M. Faheem, Cluster-based wireless sensor network framework for denial-of-service attack detection based on variable selection ensemble machine learning algorithms, *Intelligent Systems with Applications*, 2024, **22**, 200381, doi: 10.1016/j.iswa.2024.200381.

[11] G. Joseph S. G. Joseph, An adaptive deep learning framework for energy-efficient anomaly detection in IoT sensor networks, *CogNexus*, 2026, **2**, 103–119, doi: 10.63084/cognexus.v2i2.253.

[12] I. S. Khaleel, B. M. Mohammed, A. Adnan, Enhanced extreme learning machine for energy-efficient vampire attack detection in wireless sensor networks, *Journal Européen des Systèmes Automatisés*, 2024, **57**, 1649–1657, doi: 10.18280/jesa.570612.

[13] R. Kiran, K. Bhaskar, T. A. Kumar, A. D. Reddy, Optimized learning-based anomaly recognition for securing wireless sensor networks, *International Journal of Case Studies in Business, IT and Education*, 2025, 29–40, doi: 10.47992/IJCSBE.2581.6942.0369.

[14] B. Kizilkaya, E. Ever, H. Y. Yatbaz, A. Yazici, An effective forest fire detection framework using heterogeneous wireless multimedia sensor networks, *ACM Transactions on Multimedia Computing, Communications, and Applications*, 2022, **18**, 1–21, 2022, doi: 10.1145/3473037.

[15] Y. Abdullah R, B. Nisha U, Sindhu V, Palani S, A reproducible neuro-fuzzy framework for interpretable anomaly detection in wireless sensor network data streams, *Scientific Reports*, 2026, **16**, doi: 10.1038/s41598-026-68411-y.

[16] H. Rani, S. Y. Ali, M. Zubair, M. W. Riaz, R. Zubair, M. Yousif, An intelligent intrusion detection framework for wireless sensor networks using machine learning, *Spectrum of Engineering Sciences*, 2026, doi: 10.5281/zenodo.20352213.

[17] H. Sadia, S. Farhan, Y. U. Haq, R. Sana, T. Mahmood, S. A. O. Bahaj, A. R. Khan, Intrusion detection system for wireless sensor networks: A machine

learning-based approach, *IEEE Access*, 2024, **12**, 52565–52582, 2024, doi: 10.1109/ACCESS.2024.3380014.

[18] M. Saeed, R. Saeed, M. Abdelhaq, R. Alsaqour, M. Hasan, R. Mokhtar, Anomaly detection in 6G networks using machine learning methods, *Electronics*, 2023, **12**, 3300, doi: 10.3390/electronics12153300.

[19] H. Sharma, A. Haque, F. Blaabjerg, Machine learning in wireless sensor networks for smart cities: A survey," Electronics, 2021, **10**, 1012, doi: 10.3390/electronics10091012.

[20] Y. Singh T. Walingo, Smart water quality monitoring with IoT wireless sensor networks, *Sensors*, 2024, **24**, 2871, doi: 10.3390/s24092871.

[21] H. Rani, S. Younus Ali, M. Zubair, M. Waqas Riaz, R. Zubair, M. Yousif, An intelligent intrusion detection framework for wireless sensor networks using machine learning, *Spectrum of Engineering Sciences*, 2026, **4**, 1948–1965.

[22] Q. M. Zainel, A. Y. Dawod, M. F. Abdulqader, R. AlAli, M. M. AbdelRahman, A. A. Khalil, Enhanced security for wireless sensor networks through lightweight machine learning-based anomaly detection and attack classification, *Journal of Cyber Security and Risk Auditing*, 2026, **2026**, 306–335, doi: 10.63180/jcsra.thestap.2025.2.9.

Publisher Note: The views, statements, and data in all publications solely belong to the authors and contributors. GR Scholastic is not responsible for any injury resulting from the ideas, methods, or products mentioned. GR Scholastic remains neutral regarding jurisdictional claims in published maps and institutional affiliations.