Energy-aware and Context-aware Fault Detection Framework for Wireless Sensor Networks

Rosa Clavijo-López[1*], Jesús Merino Velásquez[2], Wayky Alfredo Luy Navarrete[3], Cesar Augusto Flores Tananta[4], Dorothy Luisa Meléndez Morote[5], Maria Aurora Gonzales Vigo[6] and Doris Fuster- Guillén[7]

[1*] Master, Professor, School of Management, Universidad César Vallejo, Tarapoto, Perú. rclavijol@ucvvirtual.edu.pe, Orcid: https://orcid.org/0009-0004-4168-9200

[2] Doctor, Professor, Graduate School, Universidad Nacional de Tumbes, Tumbes, Perú. jmerinov@untumbes.edu.pe, Orcid: https://orcid.org/0000-0003-3301-4487

[3] Doctor, Professor, Faculty of Business Sciences, Universidad Nacional de Tumbes, Tumbes, Perú. wluyn@untumbes.edu.pe, Orcid: https://orcid.org/0000-0003-0334-2498

[4] Doctor, Professor, Faculty of Business Sciences, Universidad César Vallejo, Tarapoto, Perú. cflorest@ucv.edu.pe, Orcid: https://orcid.org/0000-0002-9336-1483

[5] Master, Professor, School of Dentistry, Universidad de San Martin de Porres, Chiclayo, Perú. dmelendez@usmp.pe, Orcid: https://orcid.org/0000-0003-3406-9463

[6] Doctor, Professor, Graduate School, Universidad Cesar Vallejo, Chiclayo, Perú. gvigoma@ucvvirtual.edu.pe, Orcid: https://orcid.org/0000-0002-5989-6265

[7] Doctor, Professor, School of Medicine, Universidad San Juan Bautista, Lima, Perú. doris.fuster@upsjb.edu.pe, Orcid: https://orcid.org/0000-0002-7889-2243

Received: May 04, 2023; Accepted: June 18, 2023; Published: September 30, 2023

Abstract

Wireless sensor networks (WSNs) consist of many sensor nodes that are densely deployed throughout a randomized geographical area to monitor, detect, and analyze various physical phenomena. The primary obstacle encountered in WSNs pertains to the significant reliance of sensor nodes on finite battery power for wireless data transfer. Sensors as a crucial element inside Cyber-Physical Systems (CPS) renders them vulnerable to failures arising from intricate surroundings, substandard manufacturing, and the passage of time. Various anomalies can appear within WSNs, mostly attributed to defects such as hardware and software malfunctions and anomalies and assaults initiated by unauthorized individuals. These anomalies significantly impact the overall integrity and completeness of the data gathered by the networks. Therefore, it is imperative to provide a critical mechanism for the early detection of faults, even in the presence of constraints imposed by the sensor nodes. Machine Learning (ML) techniques encompass a range of approaches that may be employed to identify and diagnose sensor node faults inside a network. This paper presents a novel Energy-aware and Context-aware fault detection framework (ECFDF) that utilizes the Extra-Trees algorithm (ETA) for fault detection in WSNs. To assess the effectiveness of the suggested methodology for identifying context-aware faults (CAF), a simulated WSN scenario is created. This scenario consists of data from humidity and temperature sensors and is designed to emulate severe low-intensity

Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA), volume: 14, number: 3 (September), pp. 01-13. DOI: *10.58346/JOWUA.2023.I3.001*

*Corresponding author: Master, Professor, School of Management, Universidad César Vallejo, Tarapoto, Perú.

problems. This study examines six often-seen categories of sensor fault, including drift, hard-over/bias, spike, erratic/precision, stuck, and data loss. The ECFDF approach utilizes an Energy-Efficient Fuzzy Logic Adaptive Clustering Hierarchy (EE-FLACH) algorithm to select a Super Cluster Head (SCH) within WSNs. The SCH is responsible for achieving optimal energy consumption within the network, and this selection process facilitates the early detection of faults. The results of the simulation indicate that the ECFDF technique has superior performance in terms of Fault Detection Accuracy (FDA), False-Positive Rate (FPR), and Mean Residual Energy (MRE) when compared to other detection and classification methods.

Keywords: Wireless Sensor Networks, Super Cluster Head, EE-FLACH, Context-aware Faults, Extra-Trees Algorithm, Energy-aware Faults.

1 Introduction

The inclusion of sensors has become a crucial component in contemporary technologies such as the Internet of Things (IoT) and CPSs due to the emergence of smart applications (Ding, D., 2019). These systems acquire data from sensors deployed in the surrounding environment and transmit it to a central processing unit. The processing unit then analyzes the data to extract relevant information, which is used to determine and execute a suitable course of action. The occurrence of errors inside the sensor can result in unfavorable outcomes, as the effectiveness of these systems is strongly dependent on the accuracy and integrity of the data they gather. For example, the integrity of a product may be compromised if a sensor in an industrial CPS becomes corrupted. Similarly, the safety of miners might be jeopardized if a sensor reading in a wireless underground sensor network, which is employed to monitor underground mines, is manipulated. Hence, it is crucial to consistently verify the state of a sensor and promptly identify any deviations to prevent such occurrences (Reppa, V., 2014).

The sensor nodes can autonomously establish the sensor network's structure upon deployment within the designated area. These networks have significant interactions with the physical environment. The acquisition of data from WSNs can be achieved through the cooperative efforts of sensor nodes without human involvement. The radio frequency is the determining factor for the data transmission radius in a wireless sensor network. In the network, the communication between sensor nodes is established to ensure that each node's memory, computational power, and energy resources are constrained (Temene, N., 2022).

One of the primary challenges in the progression of WSNs is the limited power supplies compared to wired networks. This limitation results in sensor nodes expending energy during the reception, processing, and transmission of information to other nodes within the network. Consequently, the energy required for these operations is sourced from the individual batteries integrated into each node. WSNs are commonly deployed in areas characterized by high-risk conditions, rendering the recharging or replacement of batteries nearly unfeasible. The efficacy of these networks is significantly contingent upon the duration of the network's existence and the extent of its geographical reach. Using energy-conscious algorithms in developing sensor networks with extended lifespans is of utmost importance (Mishra, P.K., 2020).

Hence, the network's longevity is entirely contingent upon the power supply of its sensor nodes. This constraint gives rise to issues that serve as the foundation for several scholarly discussions concerning the lifespan of nodes and their energy usage. To conserve energy in sensor nodes, a cluster-based methodology has been devised wherein a subset of nodes can only communicate with the base station, facilitating energy preservation. Choosing an appropriate cluster head (CH) is a viable method that effectively mitigates energy usage. The utilization of clustering in WSNs has been shown to enhance the scalability of the network (Shamshirband, S., 2019). Clustering is a technique that effectively reduces the direct reliance of sensor nodes on the base station and mitigates the impact of multiple traffic loads. Additionally, it facilitates decentralizing decision-making processes for sending information at the local level.

In recent years, significant study has been conducted in the field of sensor defect detection and diagnosis, aiming to mitigate any potential negative consequences arising from sensor failure (Xu, C., 2019). In general, fault diagnostic methodologies may be classified into four categories: knowledge-based, signal-based, model-based, and hybrid procedures. Nevertheless, with the advent of cloud computing and ML, the knowledge-based (or data-driven) approach has emerged as a viable technique for detecting, diagnosing, and predicting outcomes by analyzing sensor activity using extensive datasets (Beghi, A., 2016).

This research study introduces an energy-aware and context-aware fault detection framework (ECFDF) that uses the ETA for fault detection in WSNs. To evaluate the efficacy of the proposed approach for detecting CAF, a simulated WSN scenario is generated (Akin, O., 2022). The ECFDF methodology employs the EE-FLACH algorithm to effectively choose a Super Cluster Head (SCH) in Wireless Sensor Networks (WSNs). The SCH is tasked to attain optimal energy usage throughout the network. This selection procedure serves to enable the timely identification of faults.

2 Related Works

WSNs have gained widespread use across various domains, encompassing environmental surveillance, medical care, smart cities, and industry automation. However, the deployment of these networks is frequently carried out under demanding and resource-limited settings, rendering them vulnerable to malfunctions and failures. This literature review examines the changing terrain of fault detection frameworks in WSNs, specifically emphasizing the incorporation of energy-aware and context-aware methodologies.

Saeedi Emadi and Mazinani (2018) proposed a new WSN anomaly detection method. The proposed method uses DBSCAN and SVM (Saeedi Emadi, H., 2018). This study preprocesses sensor data, clusters anomalies with DBSCAN, and classifies anomalies with SVM. Implementation involves applying the algorithm to sensor network data to detect real-time anomalies. The output values include abnormalities and their categorization. This method detects and classifies anomalies well. This method has drawbacks, like the difficulty of tweaking parameters and its sensitivity to data distribution.

According to Cheng et al. (2018), support vector regression (SVR) is used to detect defects in wireless sensor networks. The proposed method involves training SVR models with sensor data to predict sensor values (Cheng, Y., 2018). SVR models are deployed to various sensors during implementation to enable distributed fault detection. Output values include expected measurements and fault detection results. One benefit of this approach is distributed fault detection and prediction. This method has drawbacks, including model training overhead and sensor resource limits.

Jan et al. (2021) presented a machine learning-based architecture for distributed sensor-fault detection and diagnosis (Jan, S.U., 2021). The method involves collecting sensor data, preprocessing it, and using machine learning models to identify and diagnose faults. The framework is deployed into a sensor network to detect faults quickly. The output values include defect diagnoses. This method can identify and analyze errors, yielding practical and effective recommendations. However, the model's complicated implementation and resource use have drawbacks.

Jan et al. (2017) proposed a method for classifying sensor faults using SVM and statistical time-domain data. The method involves extracting features from sensor data, training SVM models, and categorizing sensor issues (Jan, S.U., 2017). The implementation process involves deploying an SVM-based classification system in a sensor network. The output values contain categorized sensor defects. One benefit of this method is fault classification accuracy. Its feature selection sensitivity and labeled training data requirement are drawbacks.

Noshad et al. (2019) proposed a random forest classifier-based WSN defect detection method. The process includes sensor data collection, random forest model training, and defect classification (Noshad, Z., 2019). Implementation involves sending the random forest classifier to sensors to find faults. Output values are categorized as sensor defects. One benefit of this method is its robust fault detection. One drawback of this method is increased model complexity and resource requirements.

Shamshirband et al. (2019) have presented FCS-MBFLEACH, an energy-efficient fault detection system for WSNs. A clustering algorithm is developed that prioritizes energy efficiency in sensor networks and defect detection systems (Shamshirband, S., 2019). Mobile sensor networks are used to implement the technology. Output values include energy-efficient network configurations and problems. The benefits of this method include energy-aware defect detection and network efficiency. However, adaptability to changing network topologies has drawbacks.

Saeed et al. (2021) introduced context-aware fault diagnosis in their study. ML techniques detect sensor faults in this system (Saeed, U., 2021). The method involves collecting contextual data, preprocessing it, and training ML models. The scheme is deployed to sensors for context-aware problem diagnostics. Output values include sensor defects and contextual information. One benefit of this approach is context-aware problem detection. However, integrating context data and deploying the model may be difficult.

Incorporating energy-aware and context-aware defect detection exhibits significant potential for enhancing the dependability and sustainability of WSNs. Utilizing this architecture presents numerous benefits, such as an extended lifespan of the network, heightened accuracy in problem detection, and increased allocation of resources.

3 Energy-aware and Context-aware Fault Detection Framework (ECFDF)

By incorporating context-awareness into energy management approaches, the framework can make educated judgments on allocating energy resources based on specific circumstances and conditions. When critical portions of the network are identified or certain periods are determined to have a greater probability of errors, the framework can assign additional energy resources towards activities such as monitoring, data gathering, and fault detection. On the other hand, in non-critical domains or under circumstances with minimal levels of risk, it is possible to optimize energy use to preserve power.

Faults are regarded as deviations from the standard functioning of sensor output. The flaws mentioned above might exhibit various characteristics, including transience, persistence, or intermittence, contingent upon the specific circumstances. Prime factors contributing to faults include network congestion, complex situations, substandard production, and sensor aging. Sensor defects are classified as deviations from typical operations within a certain trend. Mathematically, the functionality of a sensor functioning under normal conditions may be represented as Equation (1):

$$S_{fn} = F(t) + \gamma \tag{1}$$

In this equation, $F(t)$ represents the sensor output at a given time t, and γ represents the presence of noise. In an ideal scenario, the signal S_{fn} would be equivalent to the function $F(t)$. However, in practical settings, a sensor without faults would possess a certain level of related error denoted as γ. The research investigates six frequently encountered sensor faults, which may be obtained by manipulating the above equation, as elucidated in reference (Muhammed, T., 2017).

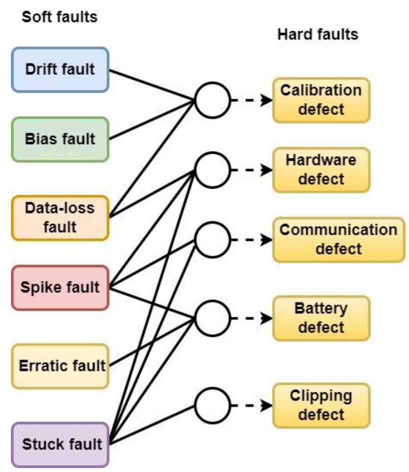

Figure 1: Various Faults Following their Respective Contextual Settings

Fig. 1 depicts the various faults following their respective contextual settings. The contexts mentioned above can be described as the internal or external surroundings of the sensor that are responsible for initiating defects. As seen in Fig. 1, it is possible for a problem to originate from either a single cause or several causes. Drift and hard-over faults are predominantly attributed to a calibration fault, but a data-loss fault can be attributed to either calibration or hardware issues. Additionally, spike faults can be attributed to several factors, such as hardware malfunctions, communication issues, or battery deficiencies. An erratic fault can be attributed to a problem in the battery, but several faults, including hardware, communication, battery, and clipping, might initiate a stuck fault.

Classification Methods

Classification, as a kind of supervised learning, is a method used to assign data findings to specific categories based on patterns and information derived from training data. The classification task is commonly categorized into three main types: binary, multi-class, and multi-label classification. Fig. 2 illustrates these classification approaches.

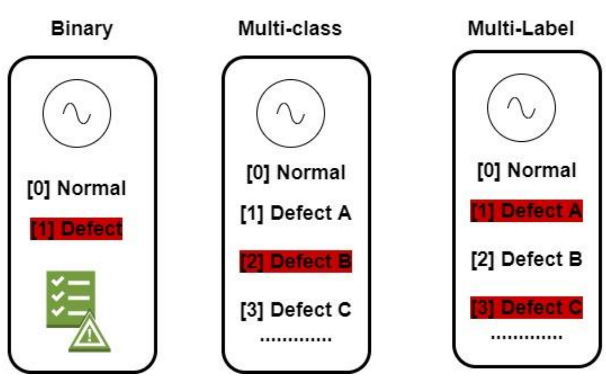

Figure 2: Fault Classification Approaches

- Binary classification is a classification task that requires categorizing data into two distinct classes. A set of data observations, sometimes called data samples, may only be categorized into one of two distinct groups. In sensor fault detection, data findings are classified into two categories: normal and abnormal.
- Multi-class classification involves the analysis of a solitary target variable representing distinct classes. This approach encompasses a set of classes that are not mutually exclusive, meaning that they can overlap or coexist. Data findings are classified into many categories according to their differences. In sensor fault diagnosis, the process involves categorizing data observations from various classes into specific target variables. These target variables often include a normal class, a drift fault class, a hard-over fault class, and other relevant classes.
- Multi-label classification is a technique used to manage many target variables inside a single class. This methodology is utilized when the data observations of a particular group simultaneously pertain to two or more variables of interest. In the case involving a trapped sensor fault class, it is seen that data concurrently pertains to numerous target variables, including hardware, communication, battery, and clipping.

Extra-Trees Algorithm

The Extra-Trees (ET) method, often referred to as Extremely Randomized Trees, is a technique that utilizes an ensemble of Decision Trees (DTs). The ensemble approach involves amalgamating numerous DTs and utilizing a fused network to make informed judgments. The RF algorithm exhibits the same flow, although with two notable distinctions. The theory behind ET involves random splitting events, whereas RF employs the best-splitting approach. Furthermore, in the case of ET, the findings are removed without replacement ($Bootstrap = False$), but in the case of RF, the insights are removed with replacement.

Moreover, the act of node splitting inside a tree refers to the procedure of converting a root node that is non-homogeneous into split/child nodes that are homogenous. The random splitting strategy employed by the Extra Trees algorithm involves dividing the root node into random child nodes. In contrast, the Random Forest algorithm utilizes a best-splitting approach that transforms the root node into homogenous child nodes. One of the primary benefits of using ETA instead of Random Forest (RF) or other ML classifiers is its ability to reduce variance and bias errors significantly. The presence of a large variance can lead to the issue of overfitting, while a high bias might induce the problem of underfitting. Additionally, the capacity of evolutionary algorithms to handle randomness enhances their computational efficiency and robustness in the presence of noisy information.

The workflow of the ETA may be elucidated through a series of four fundamental steps.

Step 1: The training set X is provided as input at the root node.

Step 2: The algorithm proceeds by selecting N samples from X in a random manner, ensuring that each sample is chosen only once and not replaced.

Step 3: A tree is constructed based on the learning samples. At each child node, a random selection of F characteristics is made, and the node is broken into arbitrary cut-points.

Step 4: The ensemble tree (ET) consolidates the results obtained from each tree by iteratively executing the second and third steps a total of T times.

In addition, it is crucial to consider some key factors while constructing the ET model. These parameters include the ensemble size k, the number of randomly selected features F, and the least number of samples required to divide a node (n_{least}).

EE-FLACH Algorithm for Selecting a SCH

The EE-FLACH method demonstrates suitability when selecting an SCH to minimize energy usage within a distributed system. To compute the fault detection rate by examining the network's nodes, applying the EE-FLACH algorithm to the proposed ECFDF approach has been chosen to provide energy-aware fault detection. The utilization of this methodology will yield significant benefits in terms of early fault identification, primarily due to its ability to minimize energy usage when compared to the MB-FLEACH, one SVM, and fuzzy one SVM techniques (Shamshirband, S., 2019). Consequently, using this approach will extend the network's operational duration. The algorithm for EE-FLACH is depicted in Algorithm 1.

Algorithm 1: EE-FLACH

Training Phase:

Input: Collection of sensor nodes, network characteristics, and limitations, such as energy thresholds, fuzzy logic membership functions and rules, and the communication model.
Output: The fuzzy logic parameters and the criterion for selecting the SCH.
Step 1: Initialization Set up the network's basic settings, such as the cluster count K, the initial fuzzy logic parameters, and the membership functions. Assign sensor nodes arbitrary or predetermined beginning energy levels. **Step 2: Cluster Formation** Based on energy levels, sensor nodes choose themselves to be CHs. Utilizing predetermined membership functions and pertinent characteristics (such as residual energy and proximity to neighbors), determine each node's fuzzy membership value for CH candidacy. As CH candidates, pick the nodes with the greatest fuzzy membership values. **Step 3: Fuzzy Logic-Based CH Selection** Use fuzzy logic principles to assess CH prospects. Consider variables like network load, residual energy, and the distance to prospective CHs. Based on these criteria, determine a weighted score for each CH candidate. Choose the CH candidate for the cluster that received the highest weighted score. **Step 4: Data Transmission and Aggregation** Each cluster's sensor nodes send information to its own CH. CHs gather and combine data from cluster members to save energy use while transmitting data. **Step 5: Energy Monitoring** Keep an eye on CHs' energy levels at all times. Determine the network's CH energy levels' weighted average. Modify the fuzzy logic's settings based on the measured performance metrics and energy levels. **Step 6: SCH Selection Criteria Determination** Based on the fuzzy logic parameters and the observed network circumstances, determine the criteria for choosing the SCH. Define the SCH membership procedures and criteria. **Output:** The results of the training phase are SCH selection criteria and fuzzy logic parameters.

Testing Phase:

Input: Set of network's sensor nodes. SCH selection criteria derived from the training phase and fuzzy logic parameters
Output: SCH selection output
Step 1: Cluster Formation Based on their energy levels, sensor nodes choose themselves as CHs. Utilizing predetermined membership functions and pertinent characteristics (such as residual energy and distance to neighbors), determine each node's fuzzy membership value for CH candidacy. As CH candidates, pick the nodes with the greatest fuzzy membership values. **Step 2: Fuzzy Logic-Based CH Selection** To assess CH candidates, use the SCH selection criteria that you got from the training step together with fuzzy logic rules. Consider variables like network load, residual energy, and the distance to prospective CHs. Based on these criteria, determine a weighted score for each CH candidate. Choose the CH candidate for the cluster that received the highest weighted score. **Step 3: SCH Election** To assess the CHs in the network, use fuzzy logic principles and the SCH selection criteria learned during the training phase. Consider elements like CH energy levels, network reach, and effectiveness of data aggregation. Create a weighted score for each CH using the information above. Choose the SCH for the whole network as the CH with the highest weighted score. **Output** The result of the testing process is the chosen SCH.

The training and testing parts of the EE-FLACH algorithm for SCH selection in WSNs are described individually in these algorithmic steps. The testing phase utilizes these parameters to choose a SCH based on real-time network circumstances. The training phase is in charge of establishing fuzzy logic parameters and SCH selection criteria.

4 Results and Discussion

The distribution of nodes within the areas is random. The performance of the suggested approach was evaluated using a 10-fold cross-validation procedure to determine the average detection accuracy for the techniques. The simulation results were conducted with the Opnet software for the specified models. Simulations were conducted with an average duration of 40 minutes for each simulation. This process was repeated ten times to ensure statistical robustness. The fault node detection was evaluated during each simulation based on the fault detection system.

The FDA metric quantifies the ratio of accurately identified defective nodes to the overall count of true fault nodes. The fault node detection methods were evaluated based on the average fault detection accuracy following the predefined scenarios. It is important to highlight that the efficient and accurate identification of faults in a network node results in minimal energy use. Consequently, it enhances the longevity of the network.

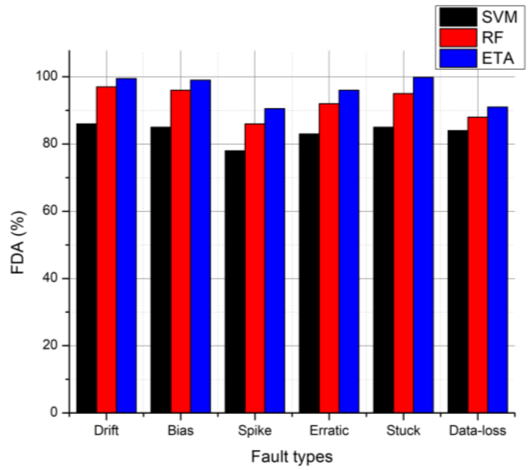

Figure 3: FDA Comparison of ML Algorithms for the Detection of Various Sensor Faults

Fig. 3 shows the FDA comparison of ML algorithms for detecting sensor faults. The findings demonstrate the efficacy of these algorithms in identifying diverse sensor malfunctions. The ETA algorithm constantly demonstrates superior performance compared to other algorithms concerning all fault types, attaining the highest FDA values. Regarding Drift defects, the ETA exhibits a notable FDA score of 99.5%. Conversely, SVM and RF attain 86% and 97% success rates correspondingly. Comparable trends are seen in various fault classifications, including Bias, Spike, Erratic, Stuck, and Data-loss. This study's results highlight the ETA algorithm's enhanced fault identification capabilities compared to SVM and RF. This suggests that the ETA algorithm holds significant potential for guaranteeing the dependability and precision of sensor fault detection in WSNs.

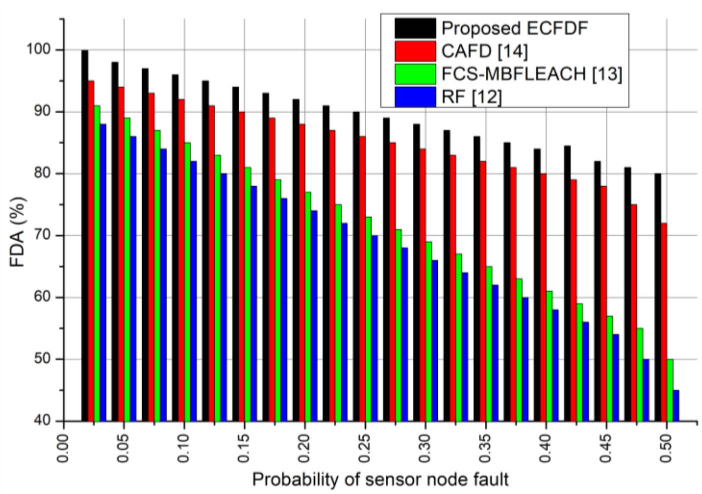

Figure 4: FDA for Various Approaches Used in Sensor Fault Detection

Fig. 4 depicts the FDA for various approaches used in sensor fault detection. This study compares four distinct techniques, specifically the Proposed ECFDF, CAFD, FCS-MBFLEACH, and RF, concerning their FDA ratings. The findings suggest that the Proposed ECFDF consistently has the best FDA across all probability levels of sensor node defects, with exceptional levels of accuracy reaching up to 99.9%. The CAFD technique demonstrates a high level of performance, as seen by FDA values reaching as high as 95%. On the other hand, it can be observed that both the FCS-MBFLEACH and RF methodologies exhibit comparatively lower FDA scores over the whole range of probabilities, with RF outperforming FCS-MBFLEACH. The results of this study emphasize the effectiveness of the proposed ECFDF method in detecting both context and energy-aware sensor faults. The findings demonstrate that the proposed ECFDF method exhibits higher accuracy than the other evaluated approaches. Consequently, it can be considered a robust option for guaranteeing the dependability and precision of fault detection in WSNs, even in diverse fault scenarios.

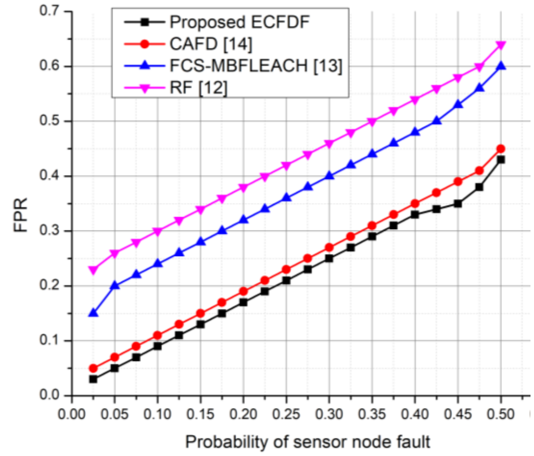

Figure 5: FPR for Various Approaches Used in Sensor Fault Detection

Fig. 5 shows the FPR for various approaches used in sensor fault detection. The FPR is a critical metric for identifying node defects within WSNs, sometimes called the false alarm rate. The criteria mentioned above are characterized by the division of the count of faultless nodes identified as faulty by the overall count of sensor nodes that are deemed faulty. The findings demonstrate that the ECFDF consistently exhibits the lowest FPR across all probability levels associated with sensor node failures. The FPR reaches values as low as 0.03 when the fault probability is 0.025. The CAFD using the FPR metric exhibits a closely trailing performance, characterized by relatively low values. On the other hand, it can be observed that the FCS-MBFLEACH and RF methodologies have elevated FPR along the whole spectrum of fault probability. The results of this study highlight the enhanced effectiveness of the proposed ECFDF approach in mitigating false alarms and enhancing the precision of fault detection, rendering it a resilient option for sensor fault detection in WSNs, particularly when the objective is to decrease false positives.

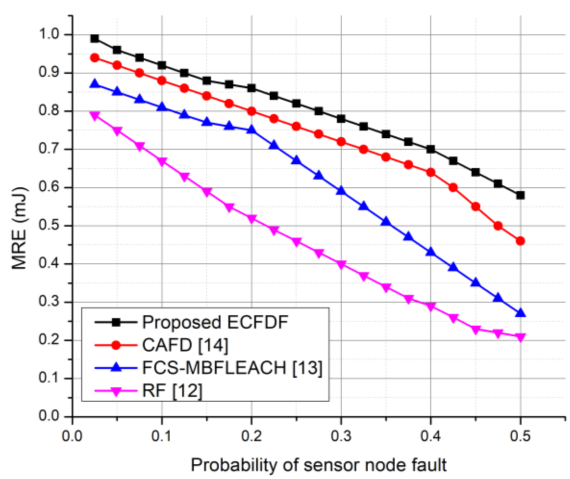

Figure 6: MRE for Various Approaches Used in Sensor Fault Detection

Fig. 6 shows the MRE for various approaches used in sensor fault detection. The findings demonstrate that the proposed ECFDF strategy consistently exhibits the highest MRE levels compared to the other methods, regardless of the fault probability. This finding suggests that using the proposed ECFDF strategy in monitoring the network of sensor nodes leads to an increase in the amount of residual energy available after fault detection and recovery. As a result, the lifetime of the network is improved. On the contrary, the remaining techniques provide diminished MRE values, indicating the likelihood of increased energy consumption by sensor nodes during fault identification and restoration, potentially influencing the network's longevity. The results of this study highlight the energy-efficient characteristics of the proposed ECFDF method. This technique effectively conserves energy in sensor nodes and prolongs the lifespan of the entire network. Consequently, it emerges as a potential option for detecting sensor faults in WSNs.

5 Conclusion

This research introduces a novel framework known as the Energy-aware and Context-aware Fault Detection Framework (ECFDF), which employs the ETA for fault detection in WSNs. To evaluate the efficacy of the proposed approach for detecting CAF, a simulated WSN scenario is generated. The presented scenario encompasses data collected from humidity and temperature sensors, intending to simulate severe low-intensity issues. This research investigates six often-seen classifications of sensor faults, including drift, hard-over/bias, spike, erratic/precision, stuck, and data loss. The ECFDF methodology employs an Energy-Efficient Fuzzy Logic Adaptive Clustering Hierarchy (EE-FLACH) algorithm to facilitate the selection of a SCH in WSNs. The SCH is tasked to attain optimal energy usage within the network. Additionally, this process of selection enables the timely identification of faults. The findings suggest that the Proposed ECFDF consistently has the best FDA and lowest FPR across all probability levels of sensor node defects, with exceptional levels of accuracy and FPR reaching up to 99.9% and 0.03, respectively.

References

[1] Akin, O., Gulmez, U.C., Sazak, O., Yagmur, O.U., & Angin, P. (2022). GreenSlice: An Energy-Efficient Secure Network Slicing Framework. *Journal of Internet Services and Information Security (JISIS), 12*(1), 57-71.

[2] Beghi, A., Brignoli, R., Cecchinato, L., Menegazzo, G., Rampazzo, M., & Simmini, F. (2016). Data-driven fault detection and diagnosis for HVAC water chillers. *Control Engineering Practice*, *53*, 79-91.

[3] Cheng, Y., Liu, Q., Wang, J., Wan, S., & Umer, T. (2018). Distributed fault detection for wireless sensor networks based on support vector regression. *Wireless Communications and Mobile Computing*, *2018*.

[4] Ding, D., Han, Q.L., Wang, Z., & Ge, X. (2019). A survey on model-based distributed control and filtering for industrial cyber-physical systems. *IEEE Transactions on Industrial Informatics*, *15*(5), 2483-2499.

[5] Jan, S.U., Lee, Y.D., & Koo, I.S. (2021). A distributed sensor-fault detection and diagnosis framework using machine learning. *Information Sciences*, *547*, 777-796.

[6] Jan, S.U., Lee, Y.D., Shin, J., & Koo, I. (2017). Sensor fault classification based on support vector machine and statistical time-domain features. *IEEE Access*, *5*, 8682-8690.

[7] Mishra, P.K., & Verma, S.K. (2020). A survey on clustering in wireless sensor network. *In IEEE 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, 1-5.

[8] Muhammed, T., & Shaikh, R.A. (2017). An analysis of fault detection strategies in wireless sensor networks. *Journal of Network and Computer Applications*, *78*, 267-287.

[9] Noshad, Z., Javaid, N., Saba, T., Wadud, Z., Saleem, M.Q., Alzahrani, M.E., & Sheta, O.E. (2019). Fault detection in wireless sensor networks through the random forest classifier. *Sensors*, *19*(7), 1-21.

[10] Reppa, V., Papadopoulos, P., Polycarpou, M.M., & Panayiotou, C.G. (2014). A distributed architecture for HVAC sensor fault detection and isolation. *IEEE Transactions on Control Systems Technology*, *23*(4), 1323-1337.

[11] Saeed, U., Lee, Y.D., Jan, S.U., & Koo, I. (2021). CAFD: context-aware fault diagnostic scheme towards sensor faults utilizing machine learning. *Sensors*, *21*(2), 1-15.

[12] Saeedi Emadi, H., & Mazinani, S.M. (2018). A novel anomaly detection algorithm using DBSCAN and SVM in wireless sensor networks. *Wireless Personal Communications*, *98*, 2025-2035.

[13] Shamshirband, S., Joloudari, J.H., GhasemiGol, M., Saadatfar, H., Mosavi, A., & Nabipour, N. (2019). FCS-MBFLEACH: Designing an energy-aware fault detection system for mobile wireless sensor networks. *Mathematics*, *8*(1), 1-24.

[14] Shamshirband, S., Joloudari, J.H., GhasemiGol, M., Saadatfar, H., Mosavi, A., & Nabipour, N. (2019). FCS-MBFLEACH: Designing an energy-aware fault detection system for mobile wireless sensor networks. *Mathematics*, *8*(1), 1-24.

[15] Temene, N., Sergiou, C., Georgiou, C., & Vassiliou, V. (2022). A survey on mobility in wireless sensor networks. *Ad Hoc Networks*, *125*.

[16] Xu, C., Zhao, S., & Liu, F. (2019). Sensor fault detection and diagnosis in the presence of outliers. *Neurocomputing*, *349*, 156-163.

Authors Biography

Rosa Clavijo-López, Master in Public Management, Bachelor in Administration, Research Professor of the School of Administration of Universidad César Vallejo.

Jesús Merino Velásquez, Doctor in Administration, Master in Economics with mention in Business Management; Professor at the Universidad Nacional de Tumbes, has developed scientific research, with publications in various international and national journals.

Wayky Alfredo Luy Navarrete, Master's degree from Universidad ESAN in Lima. Doctorate from Universidad César Vallejo. University Professor. Former Departmental Manager of ESSalud.

Cesar Augusto Flores Tananta, Finance and internal control specialist. University professor. Specialist in research methodology

Dorothy Luisa Meléndez Morote, University professor, with 17 years of teaching at undergraduate and graduate levels. I have a master's degree in Cognitive Psychopedagogy and completed studies of Doctorate in Educational Sciences.

Maria Aurora Gonzales Vigo, Systems Engineer, Universidad Señor de Sipán-Chiclayo. PhD in Public Management and Governance from Universidad Cesar Vallejo. Professor at the UCV Graduate School of Master and Doctorate; Undergraduate Professor School of Public Administration USS.

Doris Fuster- Guillén, Renacyt researcher teacher - Level III recognized by CONCYTEC, specialist in quantitative, qualitative and mixed research, PhD in Education Sciences and studying a PhD in Public Management, Master in University Teaching and research and master in information technologies and digital competences.