ANN Based Clusterhead Selection Process in WSN

Volkan GANGAL
Karadeniz Technical University
Erhan SESLI
erhansesli@ktu.edu.tr
Karadeniz Technical University
Gokce HACIOGLU
Karadeniz Technical University

Research Article

Keywords: Wireless Sensor Networks, ANN-Based WSN Clustering, Cluster Head Selection, Distributed Clustering

Posted Date: March 22nd, 2024

DOI: https://doi.org/10.21203/rs.3.rs-4109504/v1

Additional Declarations: No competing interests reported.

ANN Based Clusterhead Selection Process in WSN

Volkan GANGAL[1], Erhan SESLI[2*†], Gokce HACIOGLU[3†]

[1]Trabzon Vocational School of Higher Education, Karadeniz Technical University, Akcaabat, Trabzon, 61300, Türkiye.
[2]Of Technology Faculty, Department of Electronics and Telecommunication Engineering, Karadeniz Technical University, Trabzon, 61830, Türkiye.
[3]Department of Electrical and Electronics Engineering, Karadeniz Technical University, Trabzon, 61080, Türkiye.

*Corresponding author(s). E-mail(s): erhansesli@ktu.edu.tr;
Contributing authors: volkan.gangal@ktu.edu.tr; gokcehacioglu@ktu.edu.tr;
†These authors contributed equally to this work.

Abstract

Wireless sensor networks often face a significant challenge due to their limited energy capacity. To address this challenge, numerous studies are currently underway to explore innovative ways to optimize energy usage. This study proposes an energy-efficient cluster head selection method that is supported by artificial neural networks for cluster-based sensor networks, thereby enhancing their energy efficiency. The neural network is placed in each sensor node. The nodes decide to be the cluster head according to the threshold they set using the ANN, without the need for centralized control. According to the simulation studies conducted, the ANN trained with 5-5-5 neurons survived for 3497 rounds and successfully transmitted 96757 packets to the sink. In contrast, the LEACH algorithm could only survive for 1773 rounds and transmitted 61766 packets to the sink under the same conditions. The study confidently concluded that there was a significant and noteworthy improvement in both the network lifetime and the number of packets sent to the base station. On the other hand, it was also observed that the proposed method showed a notable improvement when compared with similar recent studies conducted with the same parameters by a factor ranging from 3.63 to 6.26.

Keywords: Wireless Sensor Networks, ANN-Based WSN Clustering, Cluster Head Selection, Distributed Clustering

1 Introduction

Wireless sensor networks (WSN) consist of multiple sensor nodes that gather information, such as temperature, humidity, images, and more, from their surroundings for different purposes, including military surveillance, patient health monitoring, and detecting natural disasters [1, 2]. Wireless sensor networks (WSNs) unavoidably face issues like network lifetime, energy, and security. Out of these, network lifetime and energy are the most critical and highly sensitive problems in a WSN [3–5]. As the energy depletion of each node can interrupt data transmission from certain regions, causing the network to become dysfunctional.

The energy consumption of nodes has a direct impact on the lifetime of a wireless sensor network. In order to increase the lifetime of the network, researchers are focusing on the reduction of the energy consumption rates during data transmission and reception. It is well known that most of the energy consumption in a node occurs during the communication phase [6, 7]. Researchers aim to develop energy saving algorithms during communication using different routing protocols. It's also known that clustering nodes has some advantages regarding energy consumption. When a node transmits its data directly to the sink, nodes far from the sink will quickly consume their energy in a WSN that is spread over a large area [8]. However, in clustered routing architecture, nodes are grouped into clusters, and a special cluster head node collects, processes, and forwards data from all sensor nodes in its cluster. In this way, member nodes save energy by sending the collected data to the cluster head node which is closer to them instead of the sink [9, 10].

The Low Energy Adaptive Clustering Hierarchy (LEACH) protocol is a well-known clustering and hierarchical approach [9–11]. Numerous works aimed at improving the performance of the LEACH algorithm have been published [9, 12–16]. Artificial neural networks (ANNs) are used in wireless sensor networks to address problems such as clustering and classification, with the aim of reducing energy consumption [17–23].

When the publications on current ANN-based energy efficiency in WSNs are evaluated in the literature, it is possible to list the main contributions of the study as follows.

- The study aimed to reduce the energy consumption of the network by determining the threshold value for cluster head selection by each node separately. This was done with the help of an Artificial Neural Network (ANN).
- In the considered scenario, node coordinates and sink are fixed, and nodes are categorised into four different categories according to their distance to the sink.
- The proposed method assumes that there is a pre-trained ANN at each node. The node decides which threshold value to use based on the output it receives by providing a set of possible threshold values as inputs to the ANN.
- The proposed method utilizes the distributed clustering mechanism of the LEACH algorithm. However, it determines the threshold values by consulting an AI model trained on past experiences.

The remainder of the paper is organized into three sections. In Section 2, we evaluate related works. In Section 3, we provide a detailed explanation of the proposed method for cluster head selection based on Artificial Neural Networks (ANN). Section 4 presents the performance results obtained from the proposed method. Finally, in Section 5, we include a discussion and draw conclusions based on the study.

2 Related Works

The notable publications in recent years addressing energy-efficient cluster head selection methods in wireless sensor networks, related to our study, are delineated in Table 1. A synopsis of the content within these publications is provided below:

A clustering algorithm was proposed by Wang et al. that utilizes an improved artificial bee colony algorithm to select cluster heads. The algorithm is based on the standard ABC algorithm but includes an efficient improved ABC algorithm. The improved ABC algorithm incorporates cluster head energy, cluster head density, cluster head location, and other similar factors to solve clustering problems in wireless sensor networks. Fuzzy C-means clustering is optimized using the improved ABC algorithm to find the most optimal clustering method. They conducted their study for three different simulation scenarios; they compared it with LEACH-C, Fitness value based Improved GWO (FIGWO), Particle Swarm Optimization (PSO), ABC-SD based studies. Wang et al. expressed by considering the simulation results that the proposed algorithm performs well in terms of energy consumption balance, energy efficiency, network life, network stability period, and network throughput [24]. Pour S.E. and Javidan R. introduced a new energy-efficient algorithm called DRE-LEACH for the selection of cluster heads (CH). This algorithm selects CHs based on several factors, including residual energy, node position, and centrality. DRE-LEACH uses a dynamic range to compute node centrality and the number of neighbors. Simulation results show that this proposed algorithm outperforms LEACH, Multi-hop Routing with LEACH (MR-LEACH), and Enhanced Multi-hop LEACH (EM-LEACH) in terms of reducing energy consumption, extending network lifetime, and enhancing network reliability [25].

Nezha et al. introduced a novel methodology, termed Energy-Aware Clustering and Efficient Cluster Head Selection (EAC-ECHS), aimed at optimizing the performance of Wireless Sensor Networks (WSNs) in terms of network life and energy efficiency. In EAC-ECHS, the sensor network is divided into inter-grid and evenly clustered grids. Moreover, for each clustered grid, cluster head selection relies on factors such as sensor residual energy, distance to neighbors, and proximity to the base station. Comparative analysis was conducted to evaluate the performance of EAC-ECHS against other protocols including LEACH, Routing G5, and OSCH G6. The findings indicate that the EAC-ECHS approach effectively meets the design objectives in terms of energy consumption, network lifetime, and packet delivery ratio [26].

Wu et al. proposed a dual cluster-head energy-efficient algorithm called DCK-LEACH, which is based on K-means and Canopy optimization. Since the K-means algorithm is sensitive to the location of initial clustering centers, this algorithm uses both the dynamic Canopy algorithm and the K-means algorithm for clustering. In cluster-head selection, DCK-LEACH adopts a hierarchical approach to minimize cluster-head overhead and balance network load. The primary cluster head is determined based on two criteria: residual energy of the node and its distance from the clustering center. Meanwhile, the vice cluster head is selected considering the node's residual energy and its distance from the base station. Simulation results demonstrate that DCK-LEACH significantly extends the lifetimes of energy-critical nodes and the overall network lifetime compared to existing protocols such as LEACH, K-LEACH, IEECHS-WSN, and RCH-LEACH [27].

In a study conducted by Gulbas G. and Cetin G., two approaches were developed to improve the lifetime of Wireless Sensor Networks (WSNs) by minimizing energy losses. They utilized the LEACH routing protocol and based their approaches on the Simulated Annealing (SA) algorithm. Both techniques, named LEACH-SA and PSCH-SA, take into consideration the residual energies and distances between the nodes when determining the Cluster Heads (CHs). The researchers conducted simulations and compared the results of both methods with LEACH. They found that PSCH-SA performed better in networks with 10 sensors, while LEACH-SA demonstrated superior performance in WSNs with 25 or more sensors [28].

In the cluster head selection process within wireless sensor networks, the utilization of an artificial neural network (ANN) is incorporated in [29]. The designated geographical area undergoes subdivision into smaller regions, with each region having a cluster head (CH) selected through the application of the ANN. Meeting a minimum energy threshold is a prerequisite for a node to qualify as a potential CH. The ANN assesses diverse node parameters to ascertain its aptness as a CH, encompassing the remaining energy level, the cumulative count of detected events, proximity to the sink, and the count of neighboring nodes.

Table 1 Comparison with recent studies

Ref.	*Authors&Year*	**Used Energy Efficiency Method**	**Compared Methods**	**Number of nodes**	**Node Deployment**	$Area[m^2]$	$E_{init}[J]$	$E_{elec}[nJ/bit]$	$E_{amp}[pJ/bit/m^2]$	$packetsize[bits]$
[24]	Wang et al. /2020	Global Artificial Bee Colony Algorithm Based on the Crossover and Tabu Search (CGTABC)	LEACH-C, FIGWHO, PSO, and ABC-SD	variable (100, 200, 300, 500)	Random and Uniform	variable(100*100, 200*200, 300*300)	0.5	50	10	4000
[25]	Pour S.E. and Javidan R. /2021	DRE-LEACH	LEACH, MR-LEACH, and EM-LEACH	100	Random and Uniform	70*70	0.5	50	10	4000
[26]	Nezha et al. /2021	Energy-Aware Clustering and Efficient Cluster Head Selection (EAC-ECHS)	LEACH, Routing G5, and OSCH G6	100	Random	100*100	0.5	50	10	4000
[27]	Wu et al. /2022	Dual Cluster-Head Energy-Efficient Algorithm (DCK-LEACH)	LEACH K-LEACH, IEECHS-WSN, and RCH-LEACH	variable (100, 200, 300)	Random	100*100	0.2	50	10	4000
[28]	Gulbas G. and Cetin G. /2023	Simulated Annealing Algorithm (SA) and LEACH, (LEACH-SA)	LEACH, and PSCH-SA	100	Random	100*100	0.5	50	10	4000
[29]	Debasis et al. /2023	ANN-Based Energy-Efcient Clustering Algorithm (EECA)	LEACH, ES-MAC, and EE-MAC	100	Random	100*100	5	Not defined	Not defined	800+160
Proposed	**Volkan et al.**	**Cascaded ANN**	**LEACH**	**100**	**Random**	**100*100**	**0.5**	**50**	**100**	**4000**

3 The Proposed Threshold Determination System

Artificial neural networks (ANN) are systems inspired by the brains of living beings, which are highly successful in solving complex problems that involve many interconnected nodes. In an ANN, weight coefficients are assigned to each link that connects the nodes to each other, and the ANN adjusts these weight coefficients using a dataset containing known inputs and outputs, learning from this process. ANN is a learning model, and its accuracy depends on the quality of the training data available to it [30]. The goal of learning in ANN is to find the optimal weights that provide accurate predictions on the given learning data. The block diagram of the cascade ANN is shown in Figure 1.

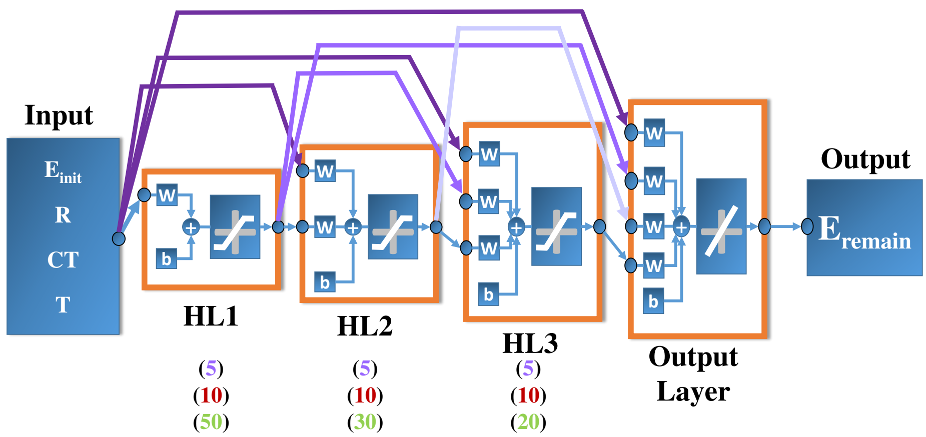

Fig. 1 Cascade-ANN block diagram used in the study

In the proposed method, nodes utilize a Cascade-ANN to determine potential threshold values. The Cascade-ANN had four inputs: the energy ratio (R), which represents the ratio of available energy to total energy of CT that the node belongs, the possible threshold value (T_i), the initial energy (E_{init}), and the category value (CT). The energy ratio (R) for a node is calculated using Equation 2, where $E_{m,\text{current}}$ and $E_{n,\text{current}}$ represent the current energy levels of the m^{th} node and the n^{th} node in the same CT, respectively. The output of the Cascade-ANN was the remaining energy (E_{remain}) of the wireless sensor node.

Furthermore, in the proposed method, Cluster Head (CH) selection takes place after 100 rounds instead of just one round. Here, one round signifies that all nodes communicate with the CH only once.

To generate the dataset, a network of 100 nodes was used, with a sink located at the center of a 100x100 meter area. The initial energy levels (E_{init}) and threshold values (T) of each node were randomly assigned. The maximum initial energy value was set to range between 0.05 and 0.5, with increments of 0.05. The $\boldsymbol{E}_{\text{initial}}$ shown in Equation 1 includes possible maximum initial energy values.

We started by selecting E_0 as the possible initial energy of the nodes, and then used the formula $E_{\text{init}} = E_0 \times \text{rand}$ to assign the E_{init} value for each node. Here, rand is a uniformly distributed random variable with values between 0 and 1. After 100 rounds of cluster head selection, the remaining energy values were stored and the nodes' energy levels were reset to $E_{\text{init}} = E_0 \times \text{rand}$. This entire process was repeated 200 times. We then replaced E_0 with E_1, and repeated the initial energy assignment process using the formula $E_{\text{init}} = E_j \times \text{rand}$, where j ranges from 0 to $g-1$. Algorithm 1 depicts the dataset creation process.

$$\boldsymbol{E}_{\text{initial}} = \begin{bmatrix} E_0 & E_1 & \ldots & E_{g-1} \end{bmatrix} \tag{1}$$

The nodes were categorized into four groups based on their distance from the sink: nodes within 15 meters were labeled as CT = 1, nodes situated between 15 and 25 meters as CT = 2, nodes positioned between 25 and 35 meters as CT = 3, and nodes located beyond 35 meters as CT = 4. Following each trial, data for variables such as E_{init}, E_{current}, T, CT, and R were collected and stored for the purpose of Cascade-ANN training. Additionally, the remaining energy of each node (E_{remain}) was recorded as the output for the Cascade-ANN training. It was assumed that all nodes, including the sink, remained stationary. In total, a dataset comprising 200,000 data points was created.

Algorithm 1 Algorithm for Creating Dataset

function CREATEDATASET($\mathbf{x}, \mathbf{y}, x_s, y_s, \boldsymbol{E}_{\text{initial}}, p_{ch}, p_{trial}$)
  $\mathbf{x}, \mathbf{y}$ - Coordinate arrays for the nodes
  Assign CT to all nodes
  $x_s = 50, y_s = 50$ ▷ Coordinates of the sink
  $p_{ch} = 100$ ▷ Cluster head election period
  $p_{trial} = 200$ ▷ Number of trials for each energy level
  for $i = 0$ to $g - 1$ **do** ▷ Number of elements of $\boldsymbol{E}_{\text{initial}}$ is g
    $E_i = \boldsymbol{E}_{\text{initial}}(i)$
    for $k = 0$ to p_{trial} **do**
      for $m = 0$ to $N - 1$ **do** ▷ N is the number of nodes
        Assign energy to m^{th} node as $E_{\text{init}} = \text{rand} \times E_i$
        Assign threshold to m^{th} node as $T = \text{rand}$
      end for
      for $q = 0$ to p_{ch} **do**
        for $m = 0$ to $N - 1$ **do**
          if rand $\leq T$ **then**
            Elect m^{th} node as Cluster Head
          end if
        end for
        Assign nearest Cluster Head to non-cluster head nodes
        Transmit and receive packets
      end for
      Calculate R in Equation 2
      Store E_{init}, T, CT, E_{remain}, and R of nodes into dataset
    end for
  end for
end function

In this study, a Cascade-ANN was used to predict the remaining energy of a wireless sensor node in a wireless sensor network. In a Cascade-ANN, there are connections from the input layer to each subsequent layer. Using the MATLAB Neural Network Toolbox, the Cascade-ANN was trained on a dataset of 200,000 samples.

$$R_m = \frac{E_{m,\text{current}}}{\sum_{n=0}^{M-1} E_{n,\text{current}}} \tag{2}$$

In the proposed method, a first-order radio model is used for the energy consumption simulations of the nodes [11]. Accordingly, the amount of energy to transmit a k-bit packet to a distance d is as in Equation 3 and the amount of energy to receive a k-bit packet is as in Equation 4.

$$E_{Tx}(k,d) = E_{\text{elec}} * k + \epsilon_{amp} * k * d^2 \tag{3}$$

$$E_{Rx}(k) = E_{\text{elec}} * k \tag{4}$$

The energy required to process a single bit in the receiver and transmitter circuits is given as E_{elec}. Meanwhile, the energy coefficient for the amplifier is represented by ϵ_{amp}.

The figure labeled as Fig. 2 illustrates the nodes belonging to the four main categories and their locations in the network. Additionally, Table 2 shows the categorization of the nodes based on their distance from the sink. The output of the Cascade-ANN is the remaining energy of the node if the node uses a possible threshold value $T_i \in \boldsymbol{Tp}$ selected from the threshold value array $\boldsymbol{Tp}$. Equation 5 shows the threshold value array, which is a row vector consisting of the threshold values T_0 through T_{m-1}.

$$\boldsymbol{T_p} = \begin{bmatrix} T_0 & T_1 & \cdots & T_{m-1} \end{bmatrix} \tag{5}$$

Since the neural network output is the E_{remain}, the index of the element in the threshold values array ($\boldsymbol{T_p}$) that corresponds to possible T_i values yielding the largest E_{remain} can be found by trying

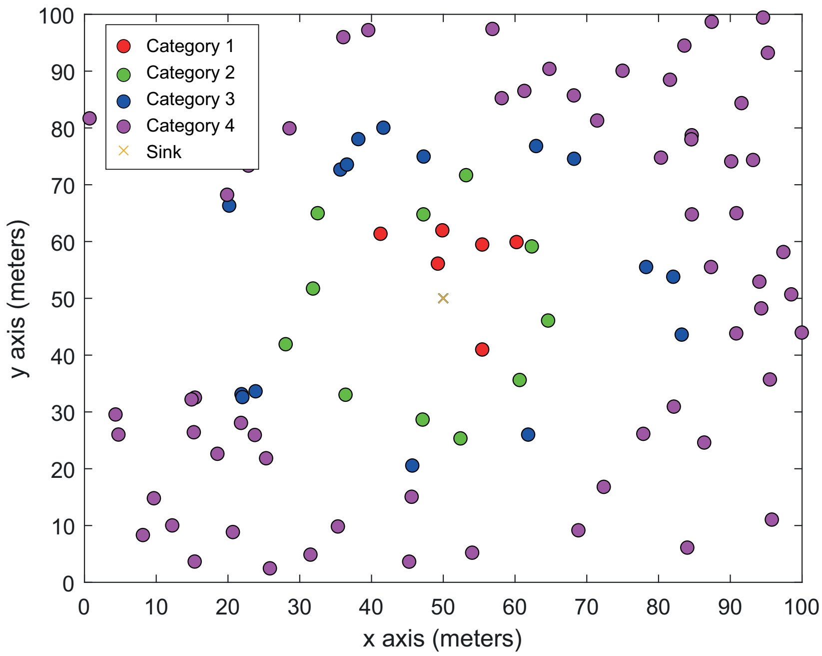

Fig. 2 Representation of the four distance categories in a 100x100 meter area

Table 2 Categories of the nodes based on distance to the Sink

Distance to the Sink (u in meters)	Category
u < 15	1
15 ≤ u < 25	2
25 ≤ u < 35	3
u ≥ 35	4

all possible threshold values one by one.

$$\mathrm{T} = \arg\max_{T_i \in \boldsymbol{T_p}} E_{\text{remain}}; i = 0, 1, \ldots, m-1 \tag{6}$$

It is presumed that each node has its own ANN and $\boldsymbol{T_p}$ saved in memory, allowing it to determine the optimal threshold value to use. After 100 rounds, each node repeats the process of utilizing its Cascade-ANN to identify the most appropriate threshold value to use.

4 Simulation Results

The simulations were performed using MATLAB. The packets had a size of 2000 bits. It was assumed that the nodes have knowledge about potential thresholds, initial energy levels, and category information for cluster head selection. The receiver and transmitter electronic circuits were assumed to use 50 nJ of energy per bit for processing ($E_{\text{elec}} = 50nJ/bit$). The amplifier energy coefficient, ϵ_{amp}, was set at $100pJ/bit/m^2$. The proposed method was simulated using the algorithm described in Algorithm 2.

Figures 3., 4. depict the training performance of the network when each hidden layer contains 5 neurons. According to Fig. 3., the network was trained for 359 epochs, resulting in the best validation performance with a mean squared error (MSE) value of 0.002454. Additionally, upon analyzing the regression plot in Fig. 4., it is evident that there exists a correlation coefficient (R) of at least 0.9178 between the data and the network's output.

We applied the proposed method to a Wireless Sensor Network (WSN), using node locations shown in Figure 2. and leveraging the trained cascade neural network. The performance of the proposed method is illustrated in Figures 5., 6., and 7.

Algorithm 2 Algorithm of the Proposed Method

```
Load the trained cascade ANN
x, y                                                    ▷ Coordinate arrays for the nodes
Assign CT to all nodes
x_s = 50, y_s = 50                                      ▷ Coordinates of the sink
p_ch = 100                                              ▷ Cluster head election period
Assign each node E_init = 0.5 joules initial energy
Load threshold value array T_p                          ▷ Equation 5
Calculate total energy of categories
Calculate ratio values for the nodes                    ▷ Equation 2
for q = 0 to N − 1 do                                   ▷ N is the number of nodes
    R = R_q                                             ▷ Ratio for node q
    for k = 0 to m − 1 do                               ▷ m is the number of possible threshold values
        T = T_k
        Use E_init, R, CT, T as input to ANN and get E_remain as output from ANN
    end for
    Find the index k_max ∈ k that gives maximum E_remain for node q
    Set threshold value of node q as T = T_{k_max}
end for
for q = 0 to N − 1 do
    if rand < T then
        Node q becomes cluster head
    else
        Node q becomes non-cluster head
    end if
end for
Assign nearest cluster head to non-cluster head nodes
while Total energy of the nodes ≥ 0 do
    for k = 0 to p_ch − 1 do
        for q = 0 to N − 1 do                           ▷ N is the number of nodes
            Transmit and receive packets
            Reduce energy for transmitting and receiving from all nodes
        end for
    end for
    Calculate ratio values for the nodes                ▷ Equation 2
    for q = 0 to N − 1 do                               ▷ N is the number of nodes
        R = R_q
        for k = 0 to m − 1 do                           ▷ m is the number of possible threshold values
            T = T_k
            Use E_init, R, CT, T as input to ANN and get E_remain from ANN
        end for
        Find the index k_max ∈ k that gives maximum E_remain for node q
        Set threshold value of node q as T = T_{k_max}
    end for
    for q = 0 to N − 1 do
        if rand < T then
            Node q becomes cluster head
        else
            Node q becomes non-cluster head
        end if
    end for
    Assign nearest cluster head to non-cluster head nodes
end while
```

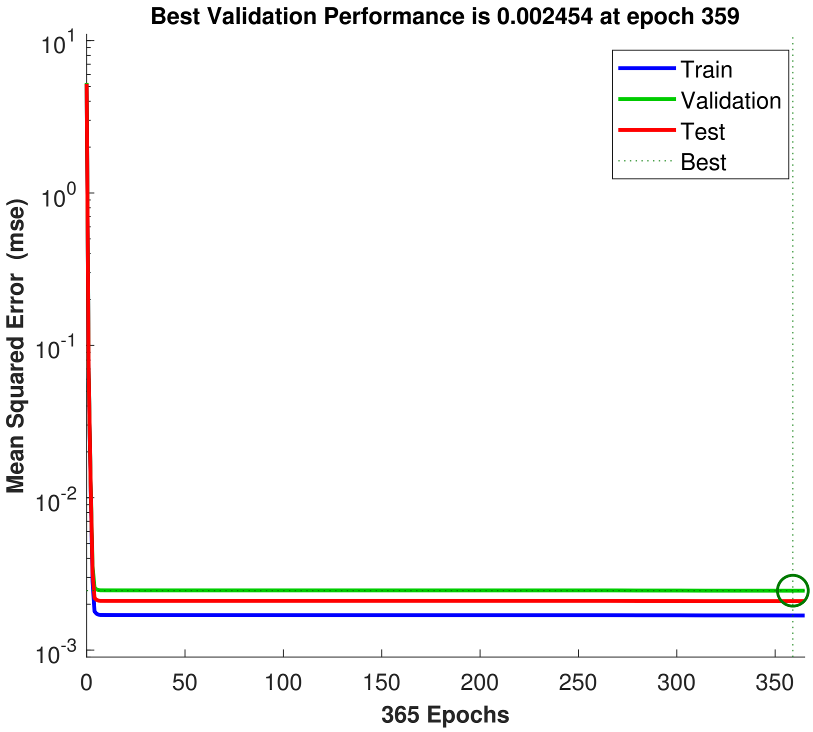

Fig. 3 Validation Performance

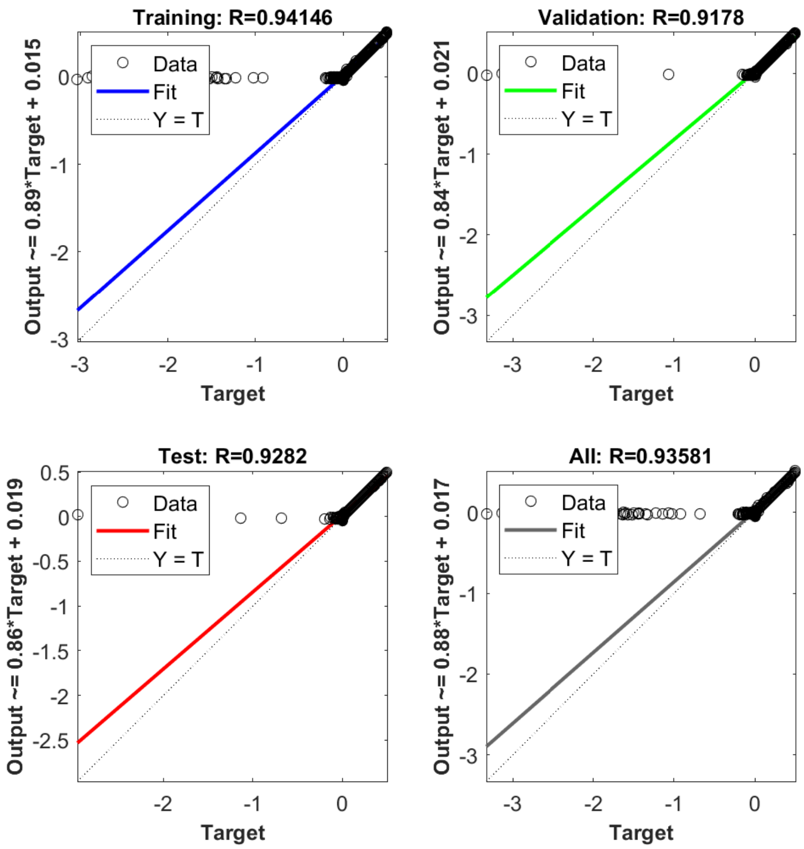

Fig. 4 Regression plot of the trained netwok

In Fig. 5., it is evident that the total energy of the proposed method surpasses that of LEACH throughout the initial rounds up to the final round. Different hidden layers for the Cascade-ANN are considered because of observing performance effects on the proposed method.

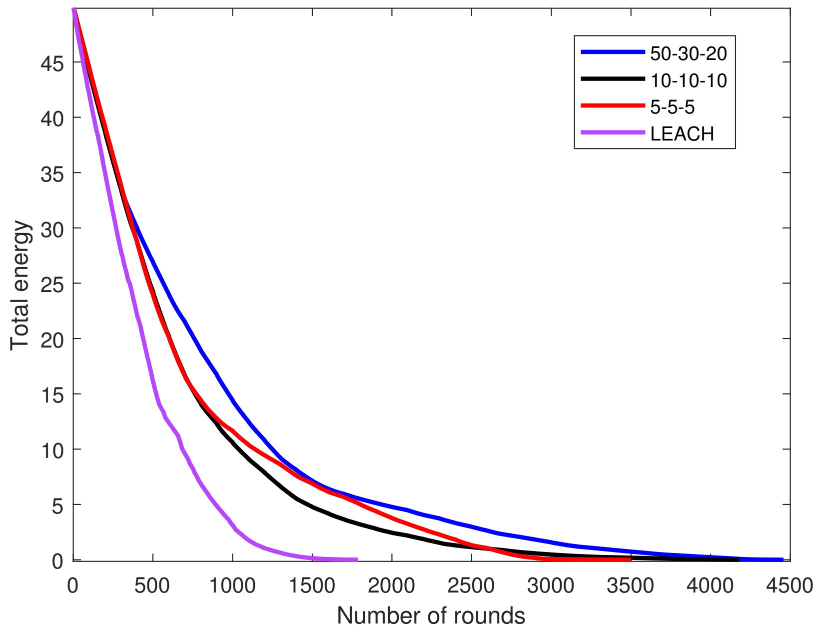

Fig. 5 Variation of the remaining energy according to the number of cycles

It can be seen from Fig. 6. that networks with a higher number of neurons in the hidden layers can keep nodes alive longer. Therefore, the goal of maximizing residual energy (E_{remain}) is better achieved by using hidden layers with more neurons. It also appears that 40% of the nodes survive around the 1500th round, regardless of the number of neurons in the hidden layers.

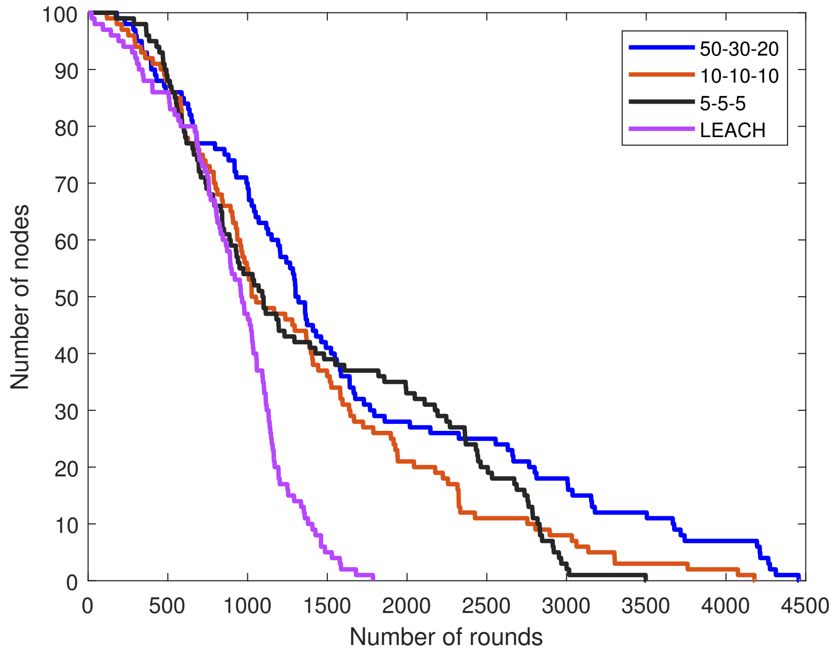

Fig. 6 Variation of the number of live nodes with respect to the number of cycles

The ANN with 50-30-20 neurons in hidden layers, survived 4457 rounds and transmitted a total of 85318 packets to the sink in Fig. 7. The ANN trained with 10-10-10 neurons survived for 4180 rounds and transmitted a total of 96420 packets to the sink. The ANN trained with 5-5-5 neurons survived 3497 rounds and transmitted a total of 96757 packets to the sink. The LEACH algorithm survived for 1773 rounds under the same conditions and transmitted a total of 61766 packets to the sink. Accordingly, the variations of the proposed method exceeded the LEACH algorithm both in terms of total packets and network lifetime. However, given that the nodes have limited processing power, it is deemed more crucial to select the scenario with three hidden layers having five neurons, as it demands less processing power.

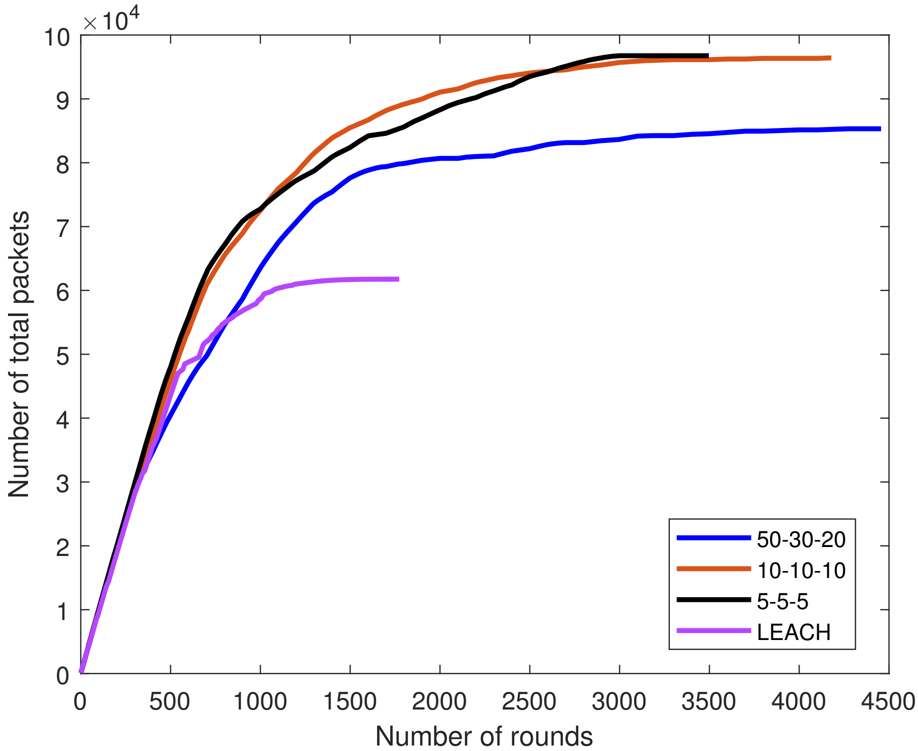

Fig. 7 Variation of the total number of packets with respect to the number of cycles

In addition, when analyzing the average thresholds, it was observed that the average threshold of the nodes in the category closest to the sink, which initially started relatively high (0.8), decreased over time (as shown in Fig. 8). This behavior can be interpreted as these nodes conserving their energy for a certain period because of their proximity to the sink. However, as their energy levels decrease, they adjust their threshold averages downward to prolong the network's lifetime.

On the other hand, nodes in Category 2 initially have the lowest average threshold values and maintain this characteristic for most of the network's operational lifespan. These lower threshold values reduce the likelihood of Category 2 nodes becoming cluster heads. It's important to note that there is a provision for increasing the threshold values of Category 2 nodes when the energy levels of Category 1 nodes fall below a certain threshold.

The average threshold for Categories 3 and 4 has not changed significantly. However, these categories are located the farthest from the sink, which means that the cluster heads in these categories consume more energy compared to the other categories. Traditionally, the solution would be to reduce the thresholds for energy efficiency. But if nodes in Categories 3 and 4 cannot find a closer cluster head, then their energy consumption while communicating with the cluster head could become even higher.

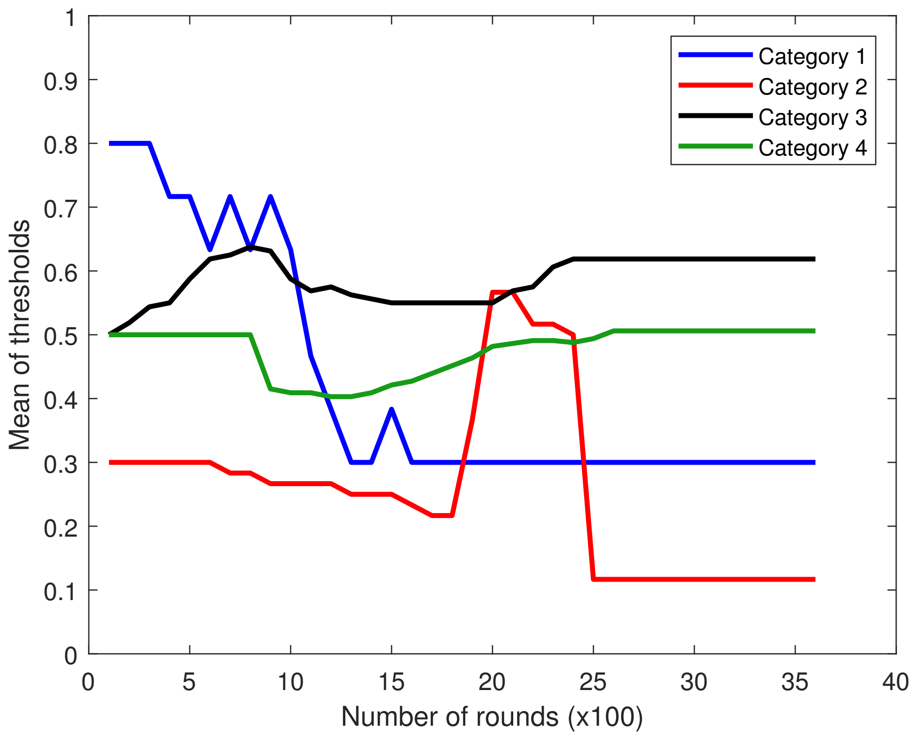

Fig. 8 Mean of thresholds

Figure 9 illustrates the energy consumption of nodes in different categories. It can be observed from Figure 9 that nodes in all categories have a lifespan exceeding 2500 rounds.

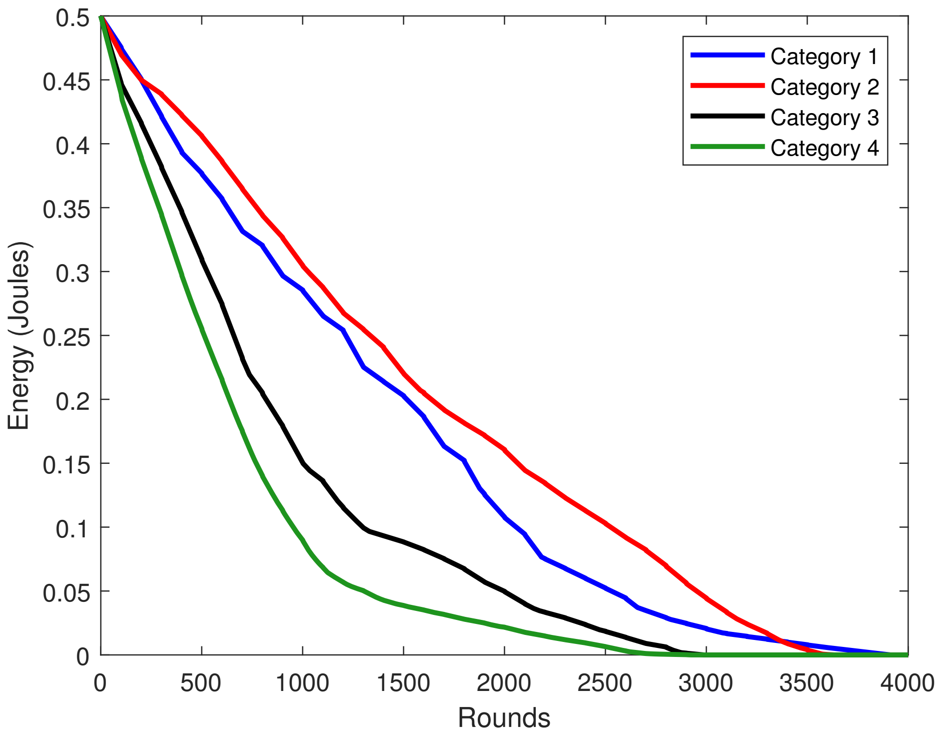

Fig. 9 Energy consumption of categories

Finally, considering the studies listed in Table 3, which have addressed related works in energy-efficient cluster head selection methods in wireless sensor networks in the recent literature, simulations were rerun with the same parameters as those in the articles. Evaluation was conducted based on the number of rounds until the last node died (LND). Furthermore, comparisons were made specifically regarding the performance achieved relative to LEACH, ensuring fairness in the assessment of results.

Table 3 Performance Analysis of the Proposed Study Between the Recent Studies

Ref. Authors	**NoN**	**ND**	*Area* $[m^2]$	E_{init} $[J]$	E_{elec} $[nJ/bit]$	E_{amp} $[pJ/bit/m^2]$	*packetsize* $[bits]$	*IRoL* [%]	*IRoLforPM* [%]
[24] Wang et al.	100	R	100*100	0.5	50	10	4000	85.9	**313.2**
[25] Pour S.E. and Javidan R.	100	R	70*70	0.5	50	10	4000	32.9	**288**
[26] Nezha et al.	100	R	100*100	0.5	50	10	4000	50	**313.2**
[27] Wu et al.	100	R	100*100	0.2	50	10	4000	69.3	**322.1**
[28] Gulbas G. and Cetin G.	100	R	100*100	0.5	50	10	4000	81.5	**313.2**
[29] Debasis et al.	100	R	100*100	5	Not defined	Not defined	800+160	31.25	**313.2**

Abbreviations: NoN ⇐ Number of Nodes, ND ⇐ Node Deployement, R ⇐ Random, IRoL ⇐ Improvement Ratio over LEACH
IRoL for PM ⇐ Improvement Ratio over LEACH for Proposed Method

In the analyses, the performance of each study calculated according to LEACH (i.e. column IRoL column in Table 3), and the proposed method was separately run based on the parameters determined in each study's simulations. Then performance of the proposed method relative to LEACH was also calculated by considering the specific parameters expressed in each study. The results obtained are indicated in the last column (i.e. column IRoL for PM in Table 3). The findings demonstrated that the proposed method outperformed similar studies by a factor ranging from 3.63 to 6.26.

5 Conclusion

We introduce a Cascade-ANN-based distributed approach for selecting Cluster Heads (CH) in wireless sensor networks. Our methodology showcases a notable enhancement in performance when compared to the conventional LEACH method, as well as some of the recent studies in this field.

It is worth noting that the proposed approach requires a database to train the Cascade-ANN. This database can be obtained either through simulations or from a physical network. Additionally, the nodes involved in this method must possess adequate processing resources to perform the necessary calculations for the ANN.

Furthermore, in future studies, it may be possible to utilize the average threshold values of categories related to the current energy levels of nodes without relying on the ANN. This could be explored by analyzing known distance spans to the sink, potentially allowing for the generalization of the proposed method without requiring knowledge of the precise node locations.

Moreover, the proposed method can be applied effectively to networks where node locations are fixed and known, such as in remote meter reading applications. In future research, we intend to implement and evaluate the proposed method in a physical network.

Declarations

- Conflict of interest: The authors have no conflicts of interest to declare that are relevant to the content of this article.
- Availability of data and materials: Not applicable
- Funding: No funding was received for conducting this study.

References

[1] Ayaz, M., Ammad-Uddin, M., Baig, I. & Aggoune, E. H. M. Wireless sensor's civil applications, prototypes, and future integration possibilities: A review. *IEEE Sensors Journal* **18**, 4–30 (2018).

[2] Kandris, D., Nakas, C., Vomvas, D. & Koulouras, G. Applications of wireless sensor networks: An up-to-date survey (2020).

[3] Ali, S. & Kumar, R. Hybrid energy efficient network using firefly algorithm, pr-pegasis and adc-ann in wsn. *Sensors International* **3**, 100154 (2022).

[4] Zivkovic, M. *et al.* Wireless sensor networks life time optimization based on the improved firefly algorithm (2020). 2020 International Wireless Communications and Mobile Computing (IWCMC), IEEE,Limassol, Cyprus, 1176–1181pp.

[5] Umbreen, S., Shehzad, D., Shafi, N., Khan, B. & Habib, U. An Energy-Efficient Mobility-Based Cluster Head Selection for Lifetime Enhancement of Wireless Sensor Networks. *IEEE Access* **8**, 207779–207793 (2020). URL https://ieeexplore.ieee.org/document/9258914/.

[6] Akyildiz, I. F. & Vuran, M. C. *Wireless sensor networks* (John Wiley & Sons, 2010).

[7] El Khediri, S. *et al.* Improved node localization using K-means clustering for Wireless Sensor Networks (2020).

[8] Sethi, D. & Bhattacharya, P. A study on energy efficient and reliable data transfer (eerdt) protocol for wban (2016). 2016 Second International Conference on Computational Intelligence & Communication Technology (CICT), IEEE, Ghaziabad, India, 254–258pp.

[9] Daanoune, I., Abdennaceur, B. & Ballouk, A. A comprehensive survey on leach-based clustering routing protocols in wireless sensor networks. *Ad Hoc Networks* **114**, 102409 (2021).

[10] Chang, L., Li, F., Niu, X. & Zhu, J. On an improved clustering algorithm based on node density for WSN routing protocol. *Cluster Computing* **25**, 3005–3017 (2022).

[11] Heinzelman, W., Chandrakasan, A. & Balakrishnan, H. Energy-efficient communication protocol for wireless microsensor networks (2000). Proceedings of the 33rd Annual Hawaii International Conference on System Sciences, IEEE, Maui, HI, USA, 10 pp. vol.2.

[12] Gamal, M., Mekky, N. E., Soliman, H. & Hikal, N. A. Enhancing the lifetime of wireless sensor networks using fuzzy logic leach technique-based particle swarm optimization. *IEEE Access* **10**, 36935–36948 (2022).

[13] Raziah, I., Yunida, Y., Away, Y., Muharar, R. & Nasaruddin, N. A new adaptive power control based on leach clustering protocol for interference management in cooperative d2d systems. *IEEE Access* **10**, 113513–113522 (2022).

[14] Sajedi, S. N., Maadani, M. & Nesari Moghadam, M. F-leach: a fuzzy-based data aggregation scheme for healthcare iot systems. *The Journal of Supercomputing* **78**, 1030–1047 (2022).

[15] Aljumaie, G. S. & Alhakami, W. A secure leach-pro protocol based on blockchain. *Sensors* **22**, 8431 (2022).

[16] Nagarajan, M., Janakiraman, N. & Balasubramanian, C. A new routing protocol for wsn using limit-based jaya sail fish optimization-based multi-objective leach protocol: an energy-efficient clustering strategy. *Wireless Networks* **28**, 2131–2153 (2022).

[17] Nehra, N., Kumar, M. & Patel, R. Neural network based energy efficient clustering and routing in wireless sensor networks (2009). 2009 First International Conference on Networks & Communications, IEEE, Chennai, India, 34–39pp.

[18] Subha, C., Malarkan, S. & Vaithinathan, K. A survey on energy efficient neural network based clustering models in wireless sensor networks (2013). 2013 International Conference on Emerging Trends in VLSI, Embedded System, Nano Electronics and Telecommunication System (ICEVENT), IEEE, Moradabad, India, 1–6pp.

[19] Osamy, W., Khedr, A. M., Salim, A., AlAli, A. I. & El-Sawy, A. A. Recent studies utilizing artificial intelligence techniques for solving data collection, aggregation and dissemination challenges in wireless sensor networks: a review. *Electronics* **11**, 313 (2022).

[20] Mishra, K. & Sharma, P. Improved cluster head selection using particle swarm optimization and neural network in wsn (2021). 2021 International Conference on Computing Sciences (ICCS),

IEEE, Phagwara, India, 13–18pp.

[21] Sharma, A. & Kansal, A. Advanced ANN Based Secured Energy Efficient Routing Protocol in WSN. *Wireless Personal Communications* **132**, 2645–2666 (2023).

[22] Kong, P. Y. Distributed Sensor Clustering Using Artificial Neural Network With Local Information. *IEEE Internet of Things Journal* **9**, 21851–21861 (2022).

[23] Revanesh, M. *et al.* Artificial neural networks-based improved Levenberg–Marquardt neural network for energy efficiency and anomaly detection in WSN. *Wireless Networks* **6** (2023). URL https://doi.org/10.1007/s11276-023-03297-6.

[24] Wang, Z., DIng, H., Li, B., Bao, L. & Yang, Z. An Energy Efficient Routing Protocol Based on Improved Artificial Bee Colony Algorithm for Wireless Sensor Networks. *IEEE Access* **8**, 133577–133596 (2020).

[25] Pour, S. E. & Javidan, R. A new energy aware cluster head selection for leach in wireless sensor networks. *IET Wireless Sensor Systems* **11**, 45–53 (2021).

[26] Nezha, E. I., Abdellah, N. & Hassan, E. A. Energy-Aware Clustering and Efficient Cluster Head Selection. *International Journal on Smart Sensing and Intelligent Systems* **14**, 1–15 (2021).

[27] Wu, M., Li, Z., Chen, J., Min, Q. & Lu, T. A Dual Cluster-Head Energy-Efficient Routing Algorithm Based on Canopy Optimization and K-Means for WSN. *Sensors* **22** (2022).

[28] Gülbaş, G. & Çetin, G. Lifetime Optimization of the LEACH Protocol in WSNs with Simulated Annealing Algorithm. *Wireless Personal Communications* **132**, 2857–2883 (2023).

[29] Debasis, K., Sharma, L. D., Bohat, V. & Bhadoria, R. S. An Energy-Efficient Clustering Algorithm for Maximizing Lifetime of Wireless Sensor Networks using Machine Learning. *Mobile Networks and Applications* **28**, 853–867 (2023).

[30] Jo, H.-S., Park, C., Lee, E., Choi, H. K. & Park, J. Path loss prediction based on machine learning techniques: Principal component analysis, artificial neural network, and gaussian process. *Sensors* **20**, 1927 (2020).