Journal of Academic Research (Applied Sciences), Vol 29, Issu 1, 2025 "75-83"

Evaluating Clustering Algorithms for Optimized Performance in Wireless Sensor Networks

Fatma Elzahra Abdulwahab Krayem[#1], Farij O. Ehtiba[#2]
[#1,2] Computer Science Department, School of Basic Science, Academy for Postgraduate Studies Misurata, Libya
[#2] Research and Consulting Center – Misurata University, Libya
[1] fatma.krayem@it.lam.edu.ly, [2] f.ehtiba@lam.edu.ly

Article information

Key words
component; Wireless Sensor Networks, WSN, Energy-Efficient, Routing Protocols, Clustering, LEACH protocol.

Received 23 **03** 2025, Accepted **07 04** 2025, Available online **09 04** 2025

Abstract

Wireless Sensor Network (WSN) has gradually become an essential technology in recent years and is employed in various fields. One of the main challenges of WSNs is their short lifespan; due to the energy consumption which is affected by communication protocols, packet data transfer, and limited battery power factors. And among these factors, many researchers are interested in energy efficiency. In this paper, a new enhanced LEACH routing technique has been proposed. The proposed technique selects cluster heads based on current energy, and it employs a root cluster head with greater current energy and a short distance to the sink to gather all data before sending them to the sink. The results exposed that the proposed technique is outperformed the standard LEACH protocol and extended the WSN network lifespan.

I. Introduction

Recently, WSN has had significant interest due to its usage in several areas such as military, medical, and environmental [1, 2]. It contains an enormous number of low-power micro-sensor nodes (SN) that are installed across a large area and are connected to at least one Base Station (BS) [3]. Every micro-sensor analyses physical or environmental parameters such as pressure, temperature, humidity, and so on [4] which then transmits the gathered data to the BS. As shown in Figure 1, a WSN consists of sensor nodes scattered at random, a BS that receives all data obtained from the environment, and a user who collects data over the internet.

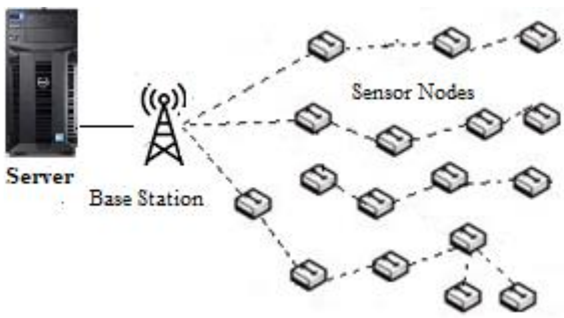

Figure 1. WSN Architecture

As shown in Figure 2, each sensor node in the network contains four units: sensing, treatment, wireless transmission, and power (usually batteries) [5], which is one of the most critical units.

f.ehtiba@lam.edu.ly

Sensor nodes in such an environment consume a lot of energy because the batteries cannot be recharged or replaced [3].

As a result, there is a growing interest in developing an energy-aware protocol to extend network lifetime [6]. Thus, in most applications where all sensor nodes are bound with energy that is related to the network's lifetime, energy consumption is the most critical element. The limited power of nodes necessitates the development of an energy-saving communication protocol [7]. Therefore, various types of research are focusing on routing protocols, which are one of the most important technologies in WSNs.

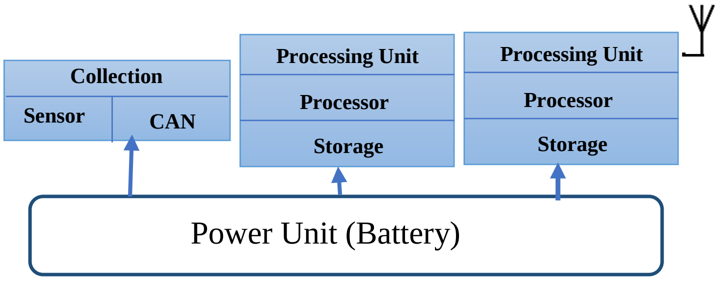

Figure 2. Sensor node units in WSN

Many routing protocols have been developed in the literature to organize data transfer from nodes to sinks [8], and the majority of them are based on the hierarchical clustering process [9]. The Low Energy Adaptive Clustering Hierarchical (LEACH) protocol [10] is a well-known and widely used as hierarchical clustering routing mechanism that will be covered in the next section [11]. Various protocols, such as DMR [12], MH-LEACH [13], MHT-LEACH [14], O-LEACH [15], and others, are suggested to extend the network lifespan and minimize energy usage in LEACH. In this work, we use the LEACH protocol to create a new enhanced LEACH protocol that reduces energy usage in WSNs.

The purpose of this study is to offer an energy-efficient, stable, residual energy, and distance-based clustering routing protocol based on the multi-hop approach. This protocol uses these parameters to reduce network energy usage and increase network lifespan in WSNs.

The rest of the paper is organized as follows: The LEACH procedure is presented in Section 2. Section 3 goes over the proposed protocol. Section 4 discusses the findings and analyses the results. Finally, the final part concludes this research.

II. LEACH (LOW ENERGY ADAPTIVE CLUSTERING HIERARCHY)

LEACH is the most popular and well-known clustering hierarchical technique intended for network energy efficiency [10]. It employs the clustering approach, which divides the network into clusters and allows the nodes to organize themselves into a hierarchical structure via the clustering process [16]. Every cycle of the LEACH algorithm has two major phases: the set-up phase and the steady phase [17].

Cluster Heads (*CHs*) are elected and clusters are created in the first phase. Every sensor node produces a random value between 0 and 1 at the start of each round [18]. This node is chosen as *CH* if its value is less than the probability function *T(n)* stated in equation (1). Otherwise, it will be a normal node [19].

$$T(n) = \begin{cases} \dfrac{p}{1 - p \times (r \, mod(1/p))}, & n \in G \\ 0 & else \end{cases} \quad (1)$$

Where *p* represents the percentage of the network's cluster count. *r* is the current epoch, and *G* is a group of nodes that have not been chosen as *CH*s in epoch *r*.

The *CHs* are chosen, and they disseminate information to the network. As a result, each normal node selects its *CH* according to the strength of the incoming signal. The second phase starts when clusters are configured and each node recognizes its TDMA time slot. The normal nodes send their collected data to their *CHs* through their TDMA time slots. The TDMA schedule is used to alter the internal access channel in each cluster to decrease inter-cluster interference [20]. Thus, each *CH* collects and compresses the data received from its members using its own, so, information from the cluster is sent straight to the BS via the *CHs* in a single hop.
LEACH is intended to reduce energy usage. It has several benefits, such as the fact that every node may, to some extent, evenly share the charge imposed at *CHs* since any sensor node that was picked as a *CH* in some epoch cannot be reelected as a *CH* yet [21]. Furthermore, using the TDMA eliminates collisions between *CHs*. However, it has certain drawbacks since it impacts the single-hop whereas *CHs* connect directly with the sink. Moreover, long-distance communication between the *CH* and the *BS* might result in higher energy usage, making LEACH unsuitable for big networks. Furthermore, because *CHs* are chosen at random, not all nodes have the opportunity to become *CHs*. As a result, if a node with lower current energy is chosen as *CH*, it will die quickly, affecting the network's life cycle.

III. THE PROPOSED PROTOCOL

The contribution of this paper is an enhanced LEACH algorithm. The purpose of the proposed protocol is to select *CH* based on node residual energy to avoid the involvement of nodes with lower energy to be *CH*. To prevent all *CHs* from communicating with the *BS*, it selects a *CH* with the highest residual energy and the shortest distance to the sink as a parent *CH*. Then, as shown in Figure 3, it leverages the multi-hop between *CHs* to reach the parent *CH*. The multi-hop approach is used to increase the distance between *CHs* and the sink. It has two major stages, which are described below.

A. Set-up phase

The random selection of *CHs* in LEACH results in a lack of chance for all nodes to become *CHs*, as well as an energy imbalance among sensor nodes. To ensure network energy balance, the proposed algorithm chooses *CH* based on the current energy. Each node computes its current energy using question (2) at each epoch. This node can play the *CH* role if its current energy exceeds the average current energy.

$$E_{cur} = E_{init} - E_{con} \quad (2)$$

$where\ E_{con} = E_{Tx} - E_{Rx}$

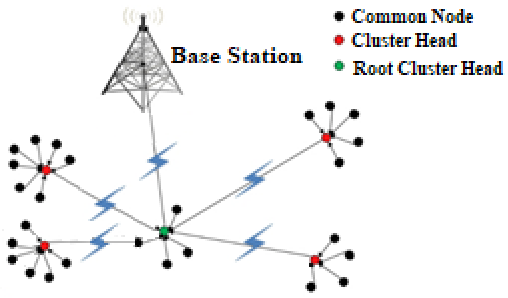

Figure 3. The Proposed Protocol Architecture

In each node, the transmitter consumes E_{Tx} to transfer *n* bits remote receiver d_i meters, and the receiver consumes E_{Rx} to receive *n* bits of a packet, and E_{init} is the initial energy provided to the nodes at the beginning.

Based on [10, 22, 23], a new threshold function has been defined as shown in question (3).

$$T(n) = \begin{cases} \dfrac{p}{1 - p_r \times \left(t \; mod\left(1/p_r\right)\right)} , n \in G \\ 0 \qquad\qquad else \end{cases} \quad , where\ PE = P_r \times \frac{E_{cur}}{E_{init}} \qquad (3)$$

Where P_r is the percentage of the number of *CHs* in the network, *t* introduces the number of the current turn, *C* is the collections of the nodes that have not yet been *CH* in the last *1/P*$_r$ turns. E_{init} represents the initial energy provides the nodes at the beginning whereas E_{cur} shows the current energy of nodes at the *t* turn.

Following the first condition, each node is assigned a random value between 0 and 1. If this value is less than the threshold function Th, this node will be elected as a *CH* on the current turn. Otherwise, it becomes a normal node. After that, *CHs* broadcast data to the network. The strength of the received signal from *CH* defines which *CH* each normal node belongs to.

B. Steady Phase

1) *TDMA (Time Division Multiple Access) schedules*

TDMA schedule is created by every *CH* after cluster formation according to the distance and the total number of nodes in each cluster [23, 24]. Each normal node has a TDMA time slot for delivering its data to its *CH*.

2) *Root cluster head selection*

In this phase, the proposed algorithm chooses the root *CH* with residual energy more than the average residual energy of *CHs* and a distance to the sink smaller than the average distance between *CHs* and the sink, as shown in questions (4) and (5).

$$E_{average} = \frac{\sum_{i=1}^{CH} E_{cur}(i)}{CH} \qquad (4)$$

$$d_{average} = \frac{\sum_{j=1}^{CH} d(j)}{CH} \qquad (5)$$

Where *d* is the distance between each *CH* and the sink, E_{cur} provides the residual energy of *CH* at the current turn.

3) Data Transmission

When all network components have learned their roles in the previous phases, the normal nodes begin gathering data from the environment and transferring it to their corresponding *CHs*. Then, using the multi-hop approach, each *CH* communicates with the nearest *CH* to reach the root *CH*. After receiving information from all *CH*s, the root *CH* aggregates it with its own and transmits it directly to the sink in a single hop, as shown in Figure 3.

V. Result and Discussion

In this section, we will utilize simulation to analyze and simulate the proposed approach's performance. The simulations are performed using the MATLAB R2016b platform. Table 1 lists the simulation parameters for the suggested model. 200 nodes have been randomly planted in an area of 200 × 200 m^2, with the sink at position (100, 250).

TABLE 1. Parameter's value

Parameter	Symbol	Value
Total number of nodes	N	200
Total number of sinks	S	1
Area	M*N	200*200m^2
Initial energy	E_{init}	1 Joule
The sink location	(x, y)	(100, 250)
The deployment of nodes	deployment	random
Energy consumed in the Radio module to transmit or receive the signal	$E_{Tx} = E_{Rx}$	50 nJ/bit
Energy consumed by the amplifier to transmit at a short distance	$^{\varepsilon}fs$	10 pJ/bit/m^2
Energy consumed by the amplifier to transmit at a longer distance	$^{\varepsilon}mp$	0.0013 pJ/bit/m^2
Data aggregation energy	EDA	5 nJ/bit
Packet size	n	500 bytes
Number of rounds	r	600

The main goal of the proposed algorithm is to increase network lifetime by selecting *CHs* based on current energy and choosing a root *CH*. All nodes begin with the same initial energy E_{init} = 1 Joule. Each sensor node sends a 4000-bit data packet to the *CH* every cycle. The simulation is performed using 10% and 20% of the total number of *CHs* in the network for both the proposed technique and LEACH.

Figure 4 (a) shows the simulation result that indicates the relative performance of the discussed protocol with parameter values from Table 1 and P_r=20%. Depending on the round, it delivers dead nodes. For P_r=10%, Figure 4 (b) compares the behavior of LEACH with the suggested approach. Figures 5(a) and (b) show the total current energy of network nodes at 20% and 10% respectively.

The two case results demonstrate the efficacy of the proposal strategy for optimizing energy usage in WSNs. Therefore, the suggested algorithm outperformed the original LEACH protocol in all circumstances, balancing overall energy usage, enhancing the stability period, and prolonging the network lifetime.

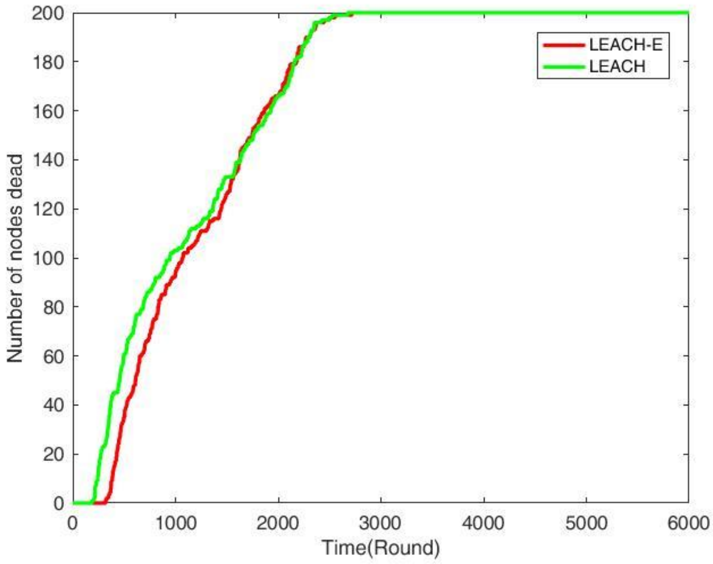

Figure 4(a): Dead nodes for 20%

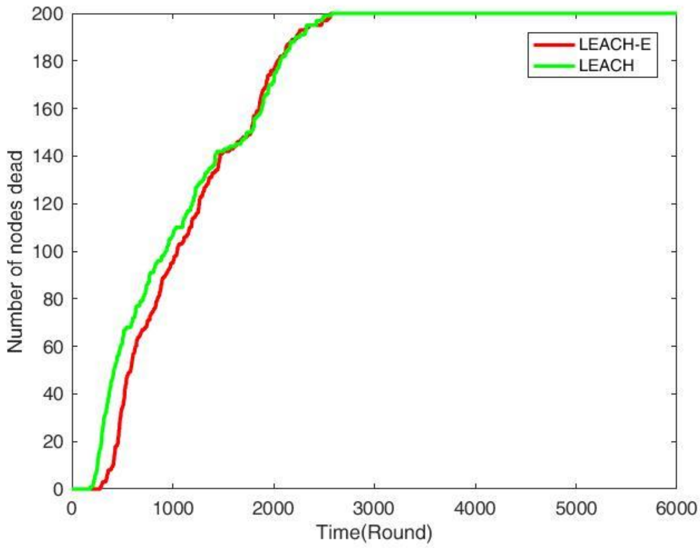

Figure 4(b): Dead nodes for 10%

Figure 4. Dead nodes in LEACH and the proposed protocol for 20% and 10%,

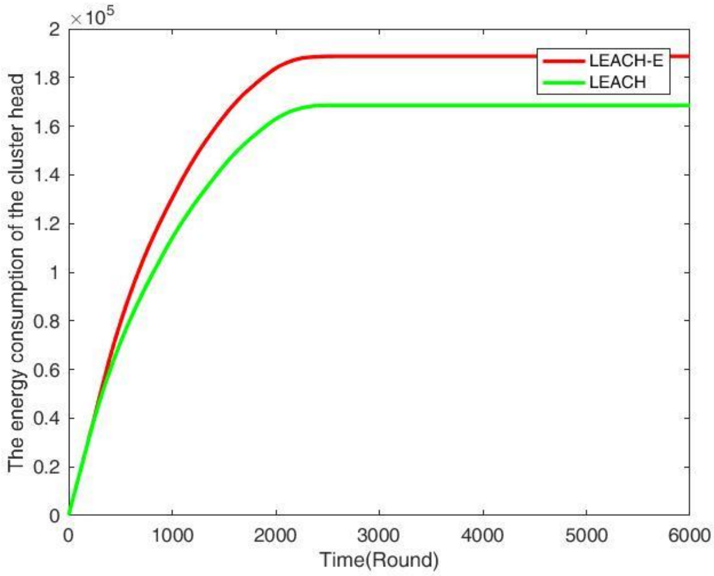

Figure 5(a): total current energy for 20%

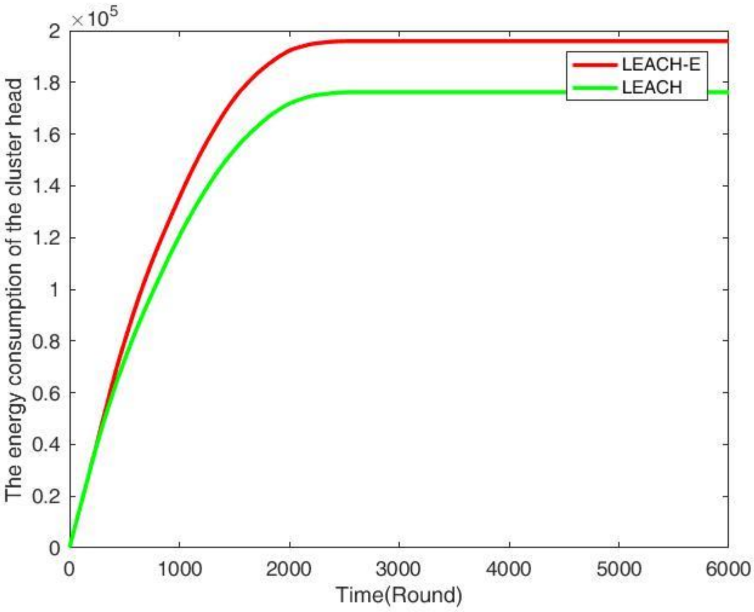

Figure 5(b): total current energy for 10%

Figure 5 shows total current energy in LEACH and the proposed protocol

V. CONCLUSION

Making an energy-efficient architecture is now the most difficult task for routing protocols in WSN. To address this issue, the sensor node must be able to operate for an extended amount of time while using less energy. In general, data transmission and reception consume the most energy in the network. Several clustering routing techniques have been proposed to tackle this difficulty. The hierarchical and clustering methods were used in this research to present a new technique for hierarchical clustering routing protocol in WSN with energy savings. The main purpose of the proposed protocol is to balance node energy usage. The second goal is to increase network longevity and stability by doing the following: (1) considering current energy in *CH* selection, (2) considering current energy and distance to sink to select a leader *CH* that sends aggregated data to sink after collecting all gathered data. The purpose of this study was to compare the suggested protocol's performance to that of the original LEACH algorithm. The simulation results verified that the proposed algorithm improves the stability period. As a result of the number of dead nodes and the total current energy of the network, it extends the network lifespan more than the original LEACH protocol. As a consequence, it reveals that the suggested method outperforms the usual LEACH approach in all scenarios addressed.

References

[1] Singh, A. and R. Kumar. Heart disease prediction using machine learning algorithms. in 2020 international conference on electrical and electronics engineering (ICE3). 2020. IEEE.

[2] Terrada, O., et al., Supervised machine learning based medical diagnosis support system for prediction of patients with heart disease. Advances in Science, Technology and Engineering Systems Journal, 2020. 5(5): p. 269-277.

[3] Yadav, S.S., et al. Automated cardiac disease diagnosis using support vector machine. in 2020 3rd International Conference on Communication System, Computing and IT Applications (CSCITA). 2020. IEEE.

[4] Gupta, A., et al., MIFH: A machine intelligence framework for heart disease diagnosis. IEEE access, 2019. 8: p. 14659-14674.

[5] Alotaibi, F.S., Implementation of machine learning model to predict heart failure disease. International Journal of Advanced Computer Science and Applications, 2019. 10(6).

[6] Kavitha, M., et al. Heart disease prediction using hybrid machine learning model. in 2021 6th international conference on inventive computation technologies (ICICT). 2021. IEEE.

[7] Chetankumar, Y.S., D.K. Singh, and S. Menaria, A review on: heart disease prediction using data mining. J Emerg Technol Innov Res (JETIR), 2019.

[8] Mohan, S., C. Thirumalai, and G. Srivastava, Effective heart disease prediction using hybrid machine learning techniques. IEEE access, 2019. 7: p. 81542-81554.

[9] El Hamdaoui, H., et al. A clinical support system for prediction of heart disease using machine learning techniques. in 2020 5th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP). 2020. IEEE.

[10] 10. Valarmathi, R. and T. Sheela, Heart disease prediction using hyper parameter optimization (HPO) tuning. Biomedical Signal Processing and Control, 2021. 70: p. 103033.

[11] Shorewala, V., Early detection of coronary heart disease using ensemble techniques. Informatics in Medicine Unlocked, 2021. 26: p. 100655.

[12] Morya, R. and S. Singh. Prediction model of heart diseases based on hybrid model. in Journal of Physics: Conference Series. 2022. IOP Publishing.

[13] Singh, P., S. Singh, and G.S. Pandi-Jain, Effective heart disease prediction system using data mining techniques. International journal of nanomedicine, 2018. 13(sup1): p. 121124.

[14] Anitha, S. and N. Sridevi, Heart disease prediction using data mining techniques. Journal of analysis and Computation, 2019.

[15] Ray, S., et al. Chi-squared based feature selection for stroke prediction using AzureML. in 2020 Intermountain Engineering, Technology and Computing (IETC). 2020. IEEE.

[16] Sekulić, A., et al., Random forest spatial interpolation. Remote Sensing, 2020. 12(10): p. 1687.

[17] Boateng, E.Y., J. Otoo, and D.A. Abaye, Basic tenets of classification algorithms Knearest-neighbor, support vector machine, random forest and neural network: a review.

Journal of Data Analysis and Information Processing, 2020. 8(4): p. 341-357.

[18] Xiong, L. and Y. Yao, Study on an adaptive thermal comfort model with K-nearestneighbors (KNN) algorithm. Building and Environment, 2021. 202: p. 108026.

[19] Das, A., Logistic regression, in Encyclopedia of Quality of Life and Well-Being Research. 2024, Springer. p. 3985-3986.

[20] Schober, P. and T.R. Vetter, Logistic regression in medical research. Anesthesia & Analgesia, 2021. 132(2): p. 365-366.

[21] Cervantes, J., et al., A comprehensive survey on support vector machine classification: Applications, challenges and trends. Neurocomputing, 2020. 408: p. 189-215.

[22] Yang, L. and A. Shami, On hyperparameter optimization of machine learning algorithms: Theory and practice. Neurocomputing, 2020. 415: p. 295-316.

[23] Daanoune, I., Baghdad, A., & Balllouk, A. (2019, October). BRE-LEACH: A new approach to extend the lifetime of wireless sensor network. In 2019 Third International Conference on Intelligent Computing in Data Sciences (ICDS) (pp. 1-6). IEEE.

[24] Al-Sadoon, M. E., Jedidi, A., & Al-Raweshidy, H. (2023). Dual-tier cluster-based routing in mobile wireless sensor network for IoT application. *IEEE Access*, *11*, 4079-4094.