



Multi-objective Unequal Optimal Clustering Algorithm for WSN Using Fuzzy Logic

Shivendra Kumar Pandey¹ · Buddha Singh¹

Received: 17 February 2023 / Accepted: 5 July 2023
© The Author(s), under exclusive licence to Springer Nature Singapore Pte Ltd 2023

Abstract

A wireless sensor network (WSN) is a network of sensors deployed in a specific area to monitor environmental characteristics. The primary objective is to collect and analyze the sensed data for deriving valuable information. Due to the limited memory, processing power, and battery capacity of sensors, all the data must be transmitted to the base station (BS) for further processing. Clustering is a commonly used method for data routing in WSNs. In this method, cluster heads (CHs) are responsible for aggregating the data within their clusters and transmitting it to the BS directly or through multi-hop transmission. However, the main challenges in this approach lie in achieving optimal CHs' selection, efficient data routing, and load balancing. An unbalanced load among CHs can lead to hot-spot problem. To address these challenges, we propose a novel clustering algorithm called the Multi-Objective Unequal Optimal Clustering Algorithm for WSN (MOUOC). This algorithm combines fuzzy logic with a linear mathematical model, considering four essential sensor parameters: residual energy, average distance of nearby nodes, the standard deviation of the distance of nearby nodes, and the distance to the BS. MOUOC selects optimal CHs with appropriate cluster size and load balancing. It solves the hot-spot problem and operates in a distributed manner, enabling high scalability. The performance of MOUOC is evaluated against four existing models across three different scenarios. The comparative analysis demonstrates that MOUOC surpasses the performance of existing models in terms of energy efficiency and network lifespan.

Keywords Energy optimization · Sensors · Multi-objective optimization · Fuzzy Logic

Introduction

A wireless sensor network (WSN) is a network of sensors deployed in a specific area to monitor environmental characteristics. Numerous industries, including the military, business, healthcare, intelligent buildings, traffic control, etc., use WSNs [1–4]. Some of the terminologies used in a standard WSN are given below:

WSN: Wireless sensors, which are distributed spatially, form a WSN [5]. The WSN consists of Cluster Heads (CHs), normal nodes, and the Base Station (BS). Here is a concise summary of these components [6].

Cluster heads: When using the clustering method of data transmission in a WSN, some sensors are chosen as CHs. The CHs obtain data from normal nodes and dispatch it to the BS after aggregating and processing.

Normal nodes: The term normal nodes or regular nodes refer to all the sensors in a cluster other than the CHs. The normal nodes transmit their data to their CHs.

Base station: The BS or sink node gathers all the sensed data of WSN for subsequent analysis. The data stored at the BS can be accessed from anywhere due to its Internet connectivity.

A WSN monitors a region to detect the occurrence of an event. Tiny sensors have limited battery life, computing power, and limited sensing capability [7–10]. Therefore, all crucial data must be gathered at the BS for further analysis. Numerous models have been proposed for efficient data

This article is part of the topical collection “Research Trends in Communication and Network Technologies” guest edited by Anshul Verma, Pradeepika Verma and Kiran Kumar Pattanaik.

✉ Shivendra Kumar Pandey
shivaiert@gmail.com
Buddha Singh
b.singh.jnu@gmail.com

¹ School of Computer and Systems Sciences, Jawaharlal Nehru University, New Mehrauli Road, New Delhi 110067, India

routing to BS. Clustering is among the most widely utilized techniques. In clustering, some normal nodes are chosen as CHs either randomly [1, 11] or based on some specific parameters, such as Residual Energy (RE), the distance to the BS, the local distance, the cluster density, etc., [12–19]. After the CHs are determined, the normal nodes join one of the CH to create the clusters. After the clusters are formed, normal nodes communicate with their CHs and forward the data. To generate a single packet, the CHs combine all of the received data with their own data [20] and transmit it to the BS either in a direct manner [1] or using multi-hop data transmission [3, 21]. However, in multi-hop transmission, the CHs closer to the BS experience higher data load, leading to early energy depletion, commonly known as the hot-spot problem [3].

To mitigate the hot-spot problem, it is essential to achieve load balancing among the CHs. Various methods have been proposed for selecting optimal CHs and balancing load among them. Fuzzy logic exhibits superior performance to other models, such as probability-based and linear mathematical models. The performance of fuzzy logic improves by increasing the input and output variables. However, as the number of input and output variables and fuzzy rules increases, the system's complexity increases exponentially. A complex fuzzy model makes the system slow, energy-consuming (due to computational work) and hard to implement on actual sensors. Therefore, instead of a complex fuzzy system, we propose a model that uses a lightweight fuzzy system along with a mathematical model for selecting optimal CHs and balancing load among them.

A lightweight fuzzy system is a fuzzy system with fewer input–output variables and fewer fuzzy rules. The lightweight fuzzy model reduces the complexity of traditional fuzzy models by minimizing the number of input–output variables and fuzzy rules. This streamlined approach brings several advantages, such as improved overall efficiency and energy savings. The simplified structure also enables faster processing speeds, making it well suited for real-time applications. The model's simplicity also facilitates easy implementation on actual sensor, making it an ideal choice for various applications. The reduced input–output variables and reduced fuzzy rules limit the performance of lightweight fuzzy logic. To achieve comparable results with reduced input–output variables and fuzzy rules, we employ a linear mathematical model in conjunction with lightweight fuzzy logic. This mathematical model helps to overcome the limitations of the lightweight fuzzy model. By combining both approaches, we aim to achieve a more energy-efficient system with higher scalability. This means that the system can efficiently handle an increasing number of nodes or sensors without sacrificing performance. The significant contributions of this paper are given below:

- In this research, we propose a novel technique for the selection of CHs by combining lightweight fuzzy logic

with a linear mathematical model. The paper introduces several parameters for sensor nodes that will be utilized by fuzzy logic and linear mathematical model. The primary goal of the algorithm is to identify the most suitable CHs along with optimal cluster size, aiming to achieve optimal energy utilization and an extended network lifetime.

- The proposed algorithm requires less computational work, making it suitable for implementation on actual sensors. It operates in a distributed manner based on local competition, which leads to improved performance and scalability. Moreover, this approach effectively reduces computational complexity and communication costs compared to centralized CH selection methods.
- We also propose an energy-efficient hierarchical routing technique to enhance the data transmission process from the CHs to the BS. This algorithm aims to minimize energy consumption while ensuring timely data delivery to the BS. The proposed algorithm uses a multi-hop approach and strategically selects intermediate nodes to relay data between the CHs and the BS. By doing so, the algorithm reduces the overall energy consumption and extends the network lifetime.
- The proposed model is evaluated by comparing it with four existing algorithms in three different scenarios. These scenarios are carefully selected to investigate the effect of the position of the BS and the type of data transmission over uniformly distributed sensors. The comparison aims to demonstrate the superiority of our proposed approach in terms of performance metrics, such as network lifetime, energy efficiency, etc.

The rest of the paper is structured as follows: The state of the art is described in the section [“Related work”](#). The assumptions and parameters used in our model are comprehensively described in the section [“Assumptions, parameters, and energy consumption model”](#). The proposed model is described in the section [“Proposed work”](#). The section [“Experiment and results”](#) compares the results with four well-known models in three different scenarios and analyses them. The section [“Conclusion and future work”](#) outlines the conclusion and its future scope. The references used in this paper are included in the last.

Related Work

To increase the energy efficiency of WSNs, numerous methods have been developed. The most often used hierarchical clustering-based technique is LEACH [1]. In LEACH, the probabilistic model is used for CHs' selection. If the probability of a node is less than a predefined threshold, it will be selected as a CH. Once a node is elected as the CH, it cannot

be considered again unless all other nodes are selected as the CH. Equation (1) is used to determine the specified threshold value of probability in LEACH

$$T(n) = \begin{cases} \frac{\rho}{1-\rho(r \bmod \frac{1}{\rho})}, & \text{if } n \in G \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

The calculated threshold value for round r is $T(n)$. The symbol ρ represents the CHs' percentage in the network. G represents the group of sensor nodes that were not chosen as CHs in previous $1/\rho$ rounds, and r represents the current round. LEACH has several limitations [22]. When choosing CHs, LEACH does not contemplate other factors like RE, sensor location, or sensor distribution. These parameters should be taken into account while selecting optimal CHs. These metrics are taken into account by several other algorithms. Similar to LEACH, EECS employs the RE of the node to choose the CHs to prolong the network lifetime [14]. The CHs are determined in HEED [23] based on probability. However, a node with higher RE is more likely to succeed. To address the issue of hot-spot, Energy Efficient Uneven Clustering (EEUC) [21] proposes the idea of unequal clustering. In EEUC, a predetermined number of nodes are initially chosen as the possible CHs. These possible CHs have different competition radii based on their distance to BS. A possible CH will be selected as final CH if it has the highest RE among other possible CHs within its competition radius. For the selection of CHs, RE is also employed in the studies cited in [24, 25]. These models do not take into account additional factors like the distance to the sink, the distribution of nodes, the average energy, the average distance between nodes, etc. Therefore, these models might choose inefficient CHs.

Evolutionary algorithms (EAs) have been widely used in WSNs for optimization and decision-making tasks due to their ability to handle complex and dynamic problems. Some EAs-based popular models are discussed here. Singh et al. [26] utilize the particle swarm optimization (PSO) approach to select CHs by incorporating distance, energy, node degree, and headcount as parameters into its fitness function. PSO is employed in several other algorithms, including [18, 27–29] by considering various parameters of sensors. For the purpose of choosing the CHs, the Firefly and Gray Wolf search-based approach is proposed in [19]. Sahoo et al. [30] proposed an improved bat algorithm for unequal clustering and efficient routing. In [31], the concept of a mobile sink is proposed for efficient data collection. It uses the PSO-GA algorithm to find the optimal position for the sink. The authors of [32] propose a clustering algorithm by combining hybrid PSO and firefly algorithm. Multi-Criteria Decision-Making (MCDM) is also employed in several algorithms. The MCDM techniques, such as AHP, ANP, TOPSIS, etc., are used in various models to select optimal CHs. In [33,

34], 7 and 12 criteria are used to determine the CHs using AHP and TOPSIS. The TOPSIS method is also used in [35–37] for efficient CHs' selection. EAs have a downside in that they demand extensive computational resources and a significant amount of data. Consequently, they can only be implemented at a centralized level (at BS), limiting the possibility of distributed CH selection algorithms. In contrast, there are alternative algorithms, such as fuzzy logic, that necessitate fewer data and computation, thus facilitating the use of distributed CH selection algorithms.

The fuzzy logic has been widely used in WSN to address issues related to imprecision and uncertainty in sensor data, energy efficiency, routing protocols, and many other aspects of WSN design and operation. In CHEF [38], fuzzy logic is used with parameters RE and local distance to calculate the CHs. In [3], the authors utilize fuzzy logic to compute the competition radius, incorporating input membership functions, such as RE, distance to BS, and density. After calculating the competition radius, the node with the highest RE is chosen as the CHs. In [2], closeness to CH and RE is used to calculate the chance of a node becoming CH using fuzzy logic. In [39], hierarchical clustering is used for the selection of CHs and super CHs using the mobility of BS, centrality, and RE. In [40], competition radius of tentative CH is calculated using the RE of the node and distance to BS. Each member node will compete with other tentative nodes to become CH in its competition radius. If node i has the highest energy among the tentative nodes within its competition radius, it will win and become CH.

In [41], a fuzzy-topsis technique based on MCDM is proposed for the election of CHs with five parameters. In [42], RE, distance, and density are used with fuzzy logic to calculate the competition radius. In the FBECS [43], the authors utilize variables, namely, RE, distance from the sink, and node density, to assess the eligibility of a node to become a CH using fuzzy logic. They establish three linguistic variables for RE, three for distance from the sink, and three for density, leading to 27 fuzzy rules. The sensing area in FBECS is divided into four equal regions based on their distance from the BS. Each region is assigned a fixed probability that determines the threshold value for CH selection. Notably, nodes in closer proximity to the BS have higher probabilities, while those in farther regions have lower probabilities. Consequently, there are a greater number of CHs with smaller cluster sizes in the vicinity of the BS, while the regions farther away have fewer CHs with larger cluster sizes. By dynamically adjusting the cluster size, FBECS aims to distribute the load among the CHs, thereby enhancing energy optimization and prolonging the lifetime of the WSN.

In the HROCF [44], the authors employ three fuzzy input variables: RE, cost, and position, to determine the most suitable CHs. They establish five linguistic variables for RE, three for cost, and three for distance, leading to 45 fuzzy

rules. By considering these variables and applying the associated rules, HROCF aims to select optimal CHs. From the literature survey, we observed that the most existing models, including FBECs [43] and HROCF [44], suffer from a common issue. While they attempt to select the best CHs by considering various parameters and employing different optimization models, they fail to address the combined aspects of load balancing among normal nodes, load balancing among CHs, and maintaining an optimal distance between the CHs. Consequently, not maintaining an optimal distance between the CHs can result in some CHs being located too close to each other sometimes, and sometimes, they may be positioned far apart. This inconsistency in CHs placement can lead to higher energy consumption during data transmission. Considering a higher number of parameters, incorporating more membership functions and fuzzy rules introduces complexities, increases energy consumption in computation work, and becomes less practical to implement on actual sensors, thereby diminishing its overall feasibility.

Hence, our objective is to develop a model that selects the optimal CHs, promotes higher scalability, and considers minimal energy consumption from CHs and normal nodes during data transmission. Furthermore, this model aims to ensure an optimal distance between CHs based on the circumstances. By achieving these goals, we aim to minimize the overall energy consumption in WSNs. For a comprehensive understanding of our proposed and existing models, we provide a summary of essential points in Table 1.

The subsequent subsection provides a detailed discussion of the assumptions, parameters, and energy consumption model used in our proposed model.

Assumptions, Parameters, and Energy Consumption Model

This section will provide the foundation for our proposed work (discussed in the section “Proposed work”). This section has been divided into four parts. In the first part, we discussed the assumptions we made while developing our model. These assumptions helped us to model the real-world problem and form a basis for our approach. In the second part, we explained how energy consumption is modeled and considered in our work, providing insights into the factors and mechanisms involved. Moving on to the third part, we explain all the parameters used in our algorithm. Finally, in the fourth part, we discussed the communication messages employed in our proposed algorithm. Each part of this section has been described sequentially, offering a detailed explanation of the respective topic.

Assumptions

We have considered several assumptions for our proposed model. These assumptions help to model the real-world problem and form a basis for our approach. These assumptions are given below:

- All the sensors are identical.
- All the sensors are deployed randomly.
- There is only one stationary sink node.
- Sensor nodes can transmit data at different energy levels.

Table 1 Comparison of different algorithms based on various criteria

Model name	CH selection	Topology	Objective	Parameter used	Scalability	Complexity
LEACH [1]	Random	Single-hop	Energy optimization	None	Low	Low
LEACH-C [45]	Random	Single-Hop	Energy optimization	RE	Low	Medium
HEED [23]	Random	Single-hop	Energy optimization	RE, node degree	Medium	Medium
TEEN [46]	Random	Multi-hop	Energy optimization	None	Low	High
PEGASIS [47]	Random	Multi-hop	Energy optimization	None (chain based routing)	Medium	High
EECS [14]	Random	Single-hop	Energy optimization	RE	Low	Low
EEUC [21]	Random	Single-hop	Energy optimization	RE	Medium	Low
EDCA [25]	Random	Single-hop	Energy optimization	RE	Medium	Low
CHEF [38]	Random	Single-hop	Energy optimization	RE, Local distance	Medium	Low
Yadav et al. [16]	Random with election	Single-Hop	Energy optimization	Position of nodes, communication range	Medium	Low
MOFCA [3]	Random with election	Multi-hop	Energy optimization	RE, distance, density	Medium	High
FBECs [43]	Random with election	Multi-hop	Energy optimization	RE, distance, density	Medium	High
HROCF [44]	Random with election	Multi-hop	Energy optimization, load balancing	RE, cost, position	Medium	High
MOUOC (proposed)	Random with election	Multi-hop	Energy optimization, load balancing	RE, distance to BS, standard deviation of local distance	High	Medium

- The receiving signal strength can be used to calculate the distance between the sender and receiver.
- All nodes begin with an equal amount of initial energy.
- After the deployment, the position of sensor nodes does not change.
- Every sensor is within the communication range of BS.
- The BS is connected to the energy source.

These assumptions will help us simulate our proposed work. We have considered that the sensor nodes consume energy in several different activities. The activities and the energy consumed in each activity are discussed below.

Energy Consumption Model

One of the crucial components of our proposed model is the calculation of energy consumption, which is achieved by utilizing the first-order radio model proposed in [1]. The sensor node's energy is depleted as it performs various tasks, such as data transmission, data reception, and data aggregation. In the transmission phase, energy is consumed by transmitter electronics and signal amplification, while in the receiving phase, energy is only consumed by receiver electronics, as shown in Fig. 1.

Three categories can be used to classify energy consumption:

- Energy consumed in transmitting the data ($E_{TX}(k, d)$).
- Energy consumed in receiving the data ($E_{RX}(k)$).
- Energy consumed in aggregation of data (E_{aggr}).

Energy consumed in transmitting the data ($E_{TX}(k, d)$): To transmit k bits of data at a distance d , the energy consumption ($E_{TX}(k, d)$) will be in transmit electronics and signal amplification. The energy consumed in transmit electronics will be $k * E_{elec}$, and the energy consumed in signal amplification will be $k * \epsilon_{fs} * d^2$ if $d < d_0$ else $k * \epsilon_{mp} * d^4$ using first-order radio model [1]. The value of d_0 is constant and

can be calculated by Eq. (6). The formula for $E_{TX}(k, d)$ is derived in Eq. (2) [1]

$$E_{TX}(k, d) = \begin{cases} k * E_{elec} + k * \epsilon_{fs} * d^2, & \text{if } d < d_0 \\ k * E_{elec} + k * \epsilon_{mp} * d^4, & \text{otherwise} \end{cases} \quad (2)$$

Energy consumed in receiving the data ($E_{RX}(k, d)$): The energy consumed in receiving k bit of data ($E_{RX}(k)$) can be calculated by Eq. (3)

$$E_{RX}(k) = k * E_{elec} \quad (3)$$

Energy consumed in aggregation of data (E_{aggr}): CHs aggregate the received data with their data and transmit the packet to the sink. If there are m normal nodes in the cluster with packet size k bits, and aggregation ratio R_{aggr} , the length of aggregated message l_{aggr} can be calculated by Eq. (4)

$$l_{aggr} = m * k * R_{aggr} + k \quad (4)$$

Energy consumption in data aggregation E_{aggr} is shown in Eq. (5)

$$E_{aggr} = l_{aggr} * E_{DA} \quad (5)$$

Here:

E_{TX} : Transmission Energy.

E_{RX} : Energy consumed in receiving k bits.

k : No. of bits in one packet.

E_{elec} : The amount of energy needed to operate electronic circuitry per bit

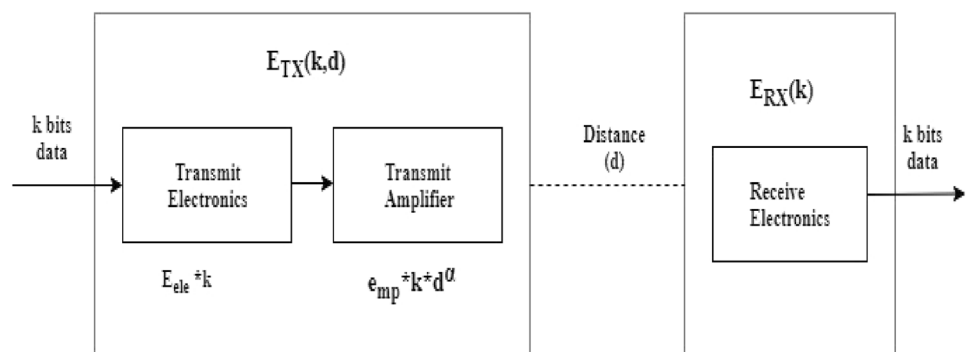
E_{DA} : The energy required to aggregate one bit of data.

ϵ_{fs} : The energy required to amplify a single bit of data using Radio Frequency (RF) when the transmission distance is less than d_0 .

ϵ_{mp} : The energy required to amplify a single bit of data using RF when the transmission distance exceeds d_0 . Here, d_0 can be calculated as in Eq. (6) [1]:

$$d_0 = \sqrt{\frac{\epsilon_{fs}}{\epsilon_{mp}}} \quad (6)$$

Fig. 1 Energy consumption model



Thus, data receiving, compilation, and transmission will all require energy from CHs. The energy consumed by a CH with m normal nodes, k bit of packet size and R_{aggr} aggregation ratio can be calculated by Eq. (7)

$$E_{\text{CH}} = E_{\text{RX}}(k) * m + E_{\text{aggr}} + E_{\text{TX}}(K, d). \quad (7)$$

Here, $E_{\text{RX}}(k)$ can be calculated by Eq. (3), E_{aggr} can be calculated by Eq. (5) and $E_{\text{TX}}(k, d)$ can be calculated by Eq. (2). The d is the distance from sender to receiver. The K is the packet size after aggregation. Normal nodes will consume energy only in data transmission to CH. If d represents the distance between a normal node and its CH, the energy consumed by the normal node in data transmission can be calculated by Eq. (2).

Utilized Parameters

We have utilized several parameters in our proposed model. These parameters will be used by our proposed model as input parameters for the computation of output parameters chance and competition radius. These parameters are explained below.

Residual energy (e): The residual energy of a node is the remaining amount of energy at the beginning of round r . In a homogeneous network, all nodes start with the same initial energy. However, due to varying amounts of energy consumption in each round, the sensor nodes may have differing amounts of energy after a few rounds. Equation (8) can be used to determine the RE of a node i at round r

$$e_i^r = E_i^o - \sum_{k=1}^{r-1} e_{ic}^k. \quad (8)$$

Here, e_i^r represents the RE on node i at the beginning of round r . The variable E_i^o represents the initial energy of the node i , and e_{ic}^k represents the amount of energy consumed by node i in round k .

Average distance of nearby nodes (d_{avg}): The parameter d_{avg} represents the average of the distance of a node with the nodes within distance R_o . Here, R_o represents the maximum value of the competition radius. The d_{avg} is a non-beneficial parameter. It means that the lower value of d_{avg} is preferable. If there are n nodes within distance R_o and $d_{\text{CH}_i}^j$ represents the distance between node j and member CH i , then the d_{avg}^i can be calculated by Eq. (9)

$$d_{\text{avg}}^i = \frac{\sum_{j=1}^n d_{\text{CH}_i}^j}{n}. \quad (9)$$

Standard deviation of the distance of nearby nodes (D_{sd}): The parameter D_{sd} denotes the standard deviation of the distances of nearby nodes that lie within a distance of R_o . If a node has a smaller value of D_{sd} , it is more likely to be chosen

as the CH. The value of parameter D_{sd} can be calculated by Eq. (10)

$$D_{\text{sd}}^i = \sqrt{\frac{\sum_{j=1}^n (d_{\text{CH}_i}^j - d_{\text{avg}}^i)^2}{n}}. \quad (10)$$

Here, d_{avg}^i represents the average distance from member CH i , and $d_{\text{CH}_i}^j$ represents the distance between sensor node j and member CH i . The communication messages used in our proposed model are discussed below.

Communication Messages

We have used three types of messages in our model. These messages are used by sensor nodes during the CH selection process. These messages are explained below.

Member_CH_Msg(ID, C, R): This message is transmitted within distance R_o by a sensor node if it is willing to participate in the CH selection process. Here, ID represents the unique integer identification number assigned to the sensor node. The variables C and R represent the chance and competition radius of the node. A node with a higher C value is more likely to be selected as CH. The competition radius represents the area in which the sensor node is competing. If a sensor node was elected CH, no other node could be elected as CH in its R .

Final_CH_Msg(ID): Once a node gets elected as CH, it will transmit the message *Final_CH_Msg(ID)* in the distance R_o , so that all the other nodes competing for CH must leave the competition for the current round.

Quit_Election_Msg(ID): If a participating node i is within distance R_o of a winning node j , then it will transmit the message *Quit_Election_Msg(ID)* to inform all the other competitor nodes within R of i that it is quitting the competition for the current round.

A detailed explanation of the proposed work and algorithms is given in the section “[Proposed work](#)”.

Proposed Work

In this section, we provide a comprehensive discussion of the CH selection, clustering, and multi-hop routing process of our proposed model, MOUOC. Our model operates in two distinct phases, each serving a specific purpose.

In the first phase, CHs are elected, and clustering is performed. Among the sensor nodes, some nodes are chosen as CHs in each round. The remaining nodes are known as normal nodes. If the distance between normal nodes and the BS is less than a defined threshold, TN_{MAX} , the normal

nodes transmit their data directly to the BS. However, if the distance exceeds TN_{MAX} , the normal nodes join the nearest CHs to form clusters. Within each cluster, the normal nodes transmit their data to their respective CHs.

In the second phase, the data routing process takes place. The CHs receive the data from the normal nodes of their clusters. They preprocess and aggregate these data into a single packet before transmitting it to the BS. If the distance between a CH and the BS is less than a defined threshold, TH_{MAX} , the CH directly transmits the data to

the BS. However, if the distance exceeds TH_{MAX} , the CH identifies the optimal intermediate node and transmits the data to it. This process continues until the data reach the BS. Once the data from all the sensors have reached the BS, the round is considered complete.

Figure 2 depicts the flowchart illustrating the steps involved in each round of MOUOC. Below, we provide a detailed elaboration on both phases of our proposed model, highlighting the specific steps and procedures undertaken in each phase.

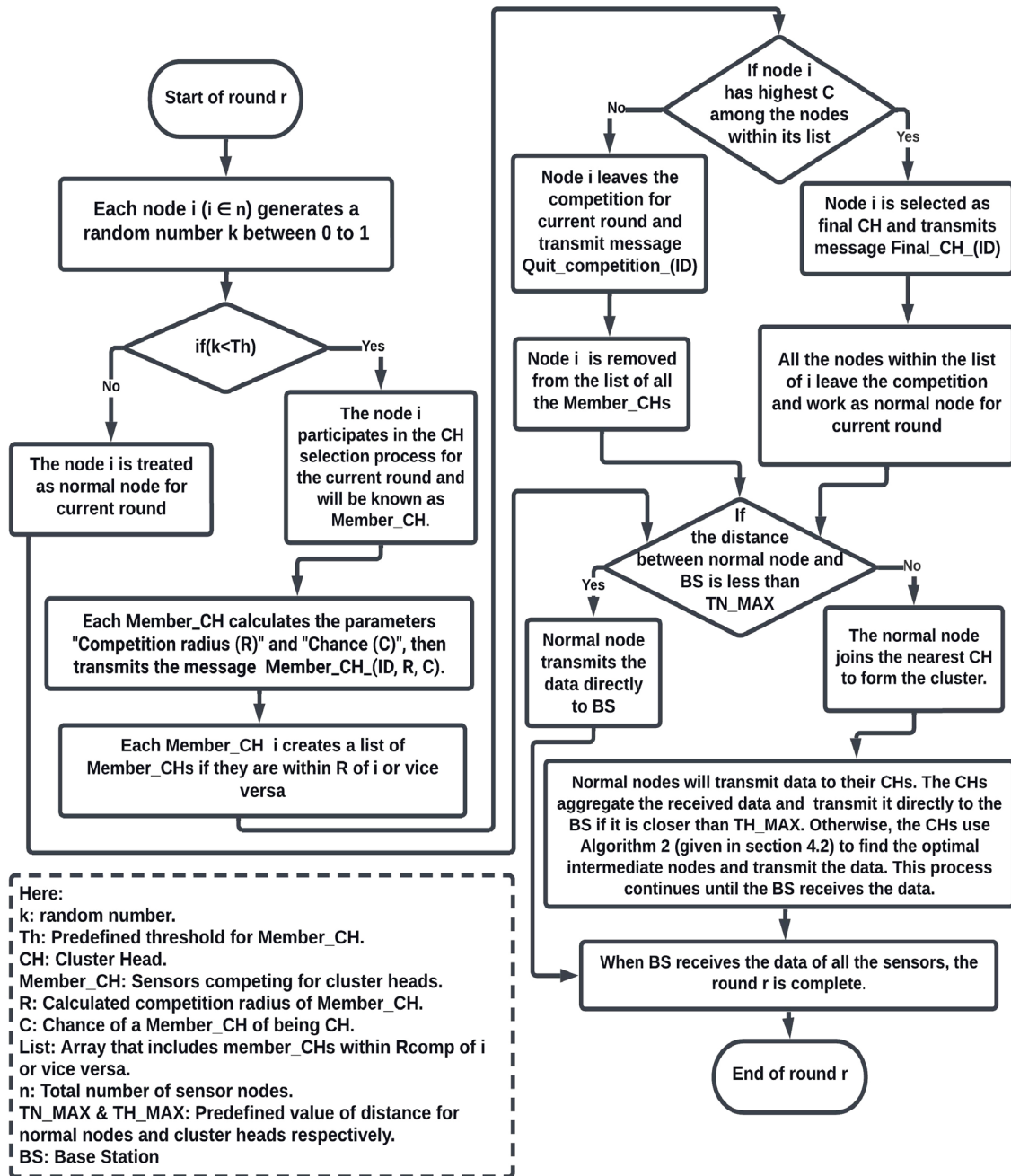


Fig. 2 Flowchart of proposed model

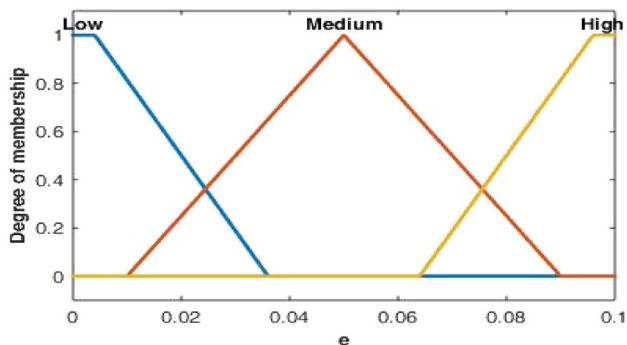


Fig. 3 The membership functions for fuzzy input variable ‘residual energy (e)’

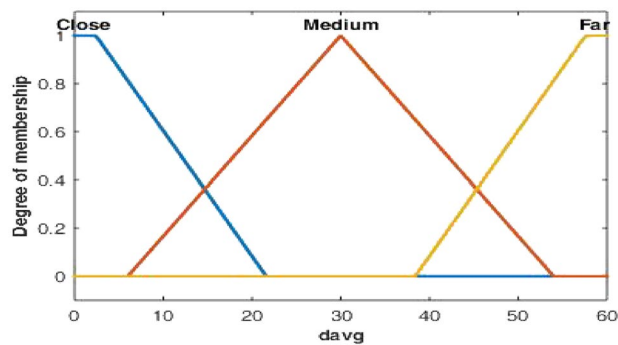


Fig. 4 The membership functions for fuzzy input variable ‘average distance of nearby nodes (d_{avg})’

Phase 1: Cluster Head Selection and Clustering

To select final CHs, initially, a fixed percentage of nodes are randomly selected as member CHs. These member CHs compete for final CHs. Each member CHs calculate two parameters for themselves: chance and competition radius. First, we will explain the process to compute these parameters and then the CH selection and clustering process. The detailed process to compute these parameters is given below.

Calculation of the Parameter Chance

To calculate the parameter chance (C), we have used fuzzy logic with three input parameters RE (e), average distance of nearby nodes (d_{avg}), and standard deviation of the distance of nearby nodes (D_{sd}). These parameters are discussed in the section “Utilized parameters” in detail. The linguistic variables for e are Low, Medium, and High, with trapezoidal membership functions for Low, High, and triangular membership function for Medium. The membership functions of e are shown in Fig. 3). The linguistic variables for d_{avg} are Close, Medium, and Far, with trapezoidal membership functions for Close, Far, and triangular membership functions for Medium. Figure 4 illustrates the membership functions for variable d_{avg} . Our third parameter is D_{sd} with two linguistic variables, Small and Large. We have used trapezoidal membership functions for both of these variables. The membership functions representing D_{sd} are shown in Fig. 5. The output membership function we have taken for our project is C . We have considered 18 linguistic variables for Chance, from C1 to C18. Here, C1 represents the slightest chance, and C18 represents the highest chance of getting selected as CH. We have used trapezoidal membership functions for C1 and C18, while triangular membership functions for C2 to C17. The membership functions for output variable C are shown in Fig. 6. We have defined 18 fuzzy rules in our fuzzy rule base. The graphical representation of these fuzzy rules

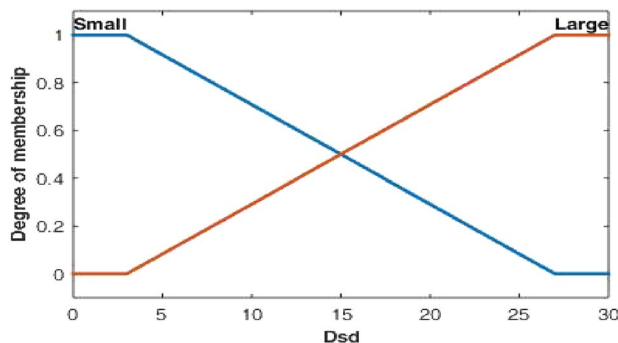


Fig. 5 The membership functions for fuzzy input variable ‘standard deviation of the distance of nearby nodes (D_{sd})’

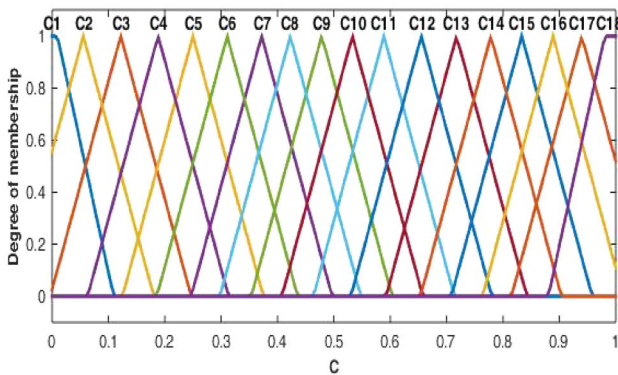
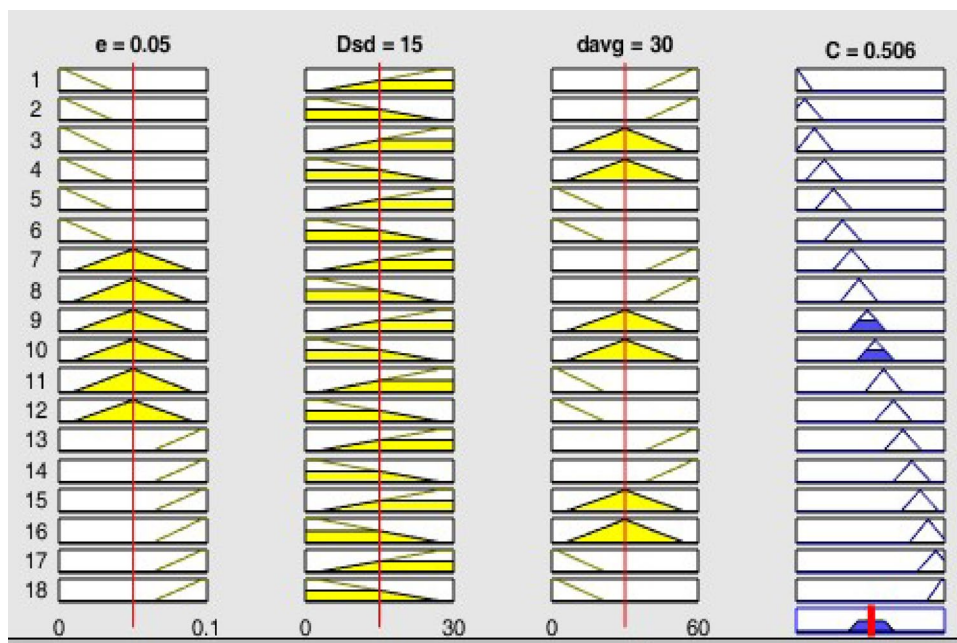


Fig. 6 The membership functions for fuzzy output variable ‘chance (C)’

is shown in Fig. 7. In this figure, the linguistic variables of all the input and output variables are represented by colored shapes. For example, in rule 1, if the e is Low, D_{sd} is Large and d_{avg} is Far, then C1 will be the output. These rules are created in such a way that a node with the lowest value e , the highest value of d_{avg} and D_{sd} has the lowest chance, while

Fig. 7 Graphical representation of proposed fuzzy rules



a node with the highest e , lowest d_{avg} and D_{sd} has highest chance of getting selected as CH.

The fuzzy model with all the inputs and output variables of our proposed work is shown in Fig. 8. We have three input membership functions which are already discussed in the section “Utilized parameters”. The fuzzy interface takes crisp input values of these input parameters, and with the set of defined rules, it generates crisp output C . For fuzzification, we have used the Mamdani method [48]. This method has been used in several past algorithms like [2, 39, 42]. To get crisp output by defuzzification, the Center of Area (COA) method is used. In the COA method, the defuzzified value χ can be calculated by Eq. (11)

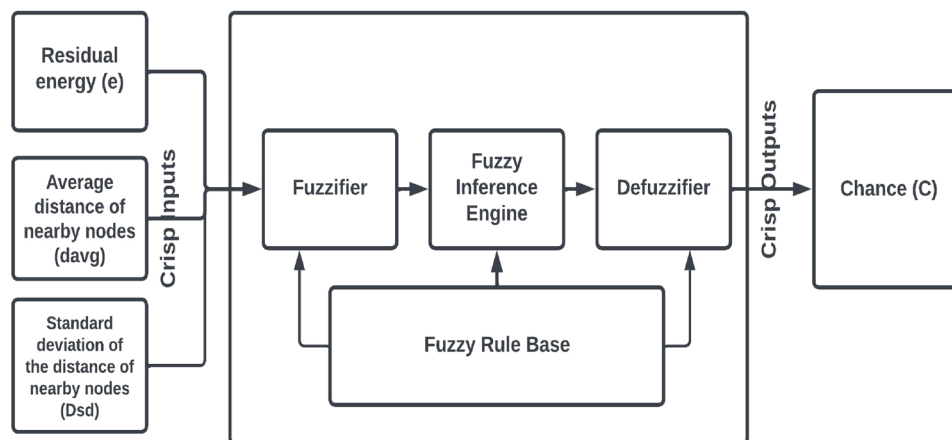
$$\chi = \frac{\int x\mu_A(x)dx}{\int \mu_A(x)dx} \tag{11}$$

The parameter C , along with another parameter R , will be used in our algorithm for the selection of final CHs. The method for computation of parameter R is discussed below.

Calculation of the Parameter Competition Radius

The competition radius (R) is calculated for all these member nodes based on their distance to sink, their RE, and the minimum and maximum distance of the sensors from BS. We assume that after the deployment of sensors, the BS transmits a hello packet to all the nodes in the network. All

Fig. 8 Input and output variables of fuzzy set



the sensor nodes calculate their distance with BS based on the received signal strength. We have already assumed that the sensors may calculate the distance based on the received signal strength. After the calculation of the distance, a small data packet with sensor ID and distance information is flooded toward BS. The BS receives all the information, calculates d_{\max} , d_{\min} , and shares it with all the sensor nodes. This process is done only once after the deployment of the sensors, and only a tiny amount of energy from the sensors is consumed in this process. After each round, if some nodes are dead, BS recalculates the value of d_{\max} , d_{\min} and shares it with all the sensor nodes again. If some new nodes are added in WSN or the position of some nodes is changed due to any reason, or any node finds its distance to BS out of the range (d_{\min}, d_{\max}) , then we have to recalculate the value of d_{\max} and d_{\min} . After getting the value of all the parameters, each node computes the value of its R using Eq. (12). A similar method to compute R is also proposed in [21], but it does not consider the RE of the nodes. The formula for the calculation of R is given in Eq. (12)

$$S(i).R = \left\{ 1 - c_1 * \frac{(d_{\max} - d(S_i, BS))}{(d_{\max} - d_{\min})} - c_2 * \frac{(E_0 - S(i).e)}{(E_0)} \right\} * R_0; \quad (12)$$

where

- R_0 = maximum competition radius.
- c_1, c_2 = predefined constant.
- d_{\max} = distance of the furthest node from sink.
- d_{\min} = distance of the closest node from sink.
- E^0 = initial energy of the node.
- $S(i).e$ = RE of node i .
- $d(S_i, BS)$ = distance of node i from sink node.
- $S(i).R$ = competition radius of node i .

It can be observed from Eq. (12) that a node with higher RE has higher R . A node's R is lower if it is nearer the BS, whereas it is higher if it is farther away from BS. To prevent hot-spot situations and prolong the lifespan of the network, we adjusted the value of R to distribute the load

evenly across the CHs. Here, the predefined constants c_1 and c_2 are used to decide the range of R . It can be observed from Eq. (12) that the range of R will be from $(1-c_1-c_2)*R_0$ to R_0 . We have tuned our values of c_1 and c_2 manually to get the optimal results. The parameter R , along with another parameter C , will be used in our proposed algorithm for the computation of CHs. Our proposed algorithm for the selection of the final CHs is given in the subsequent subsection.

Algorithm for Cluster Head Selection and Clustering

We assume that the WSN consists of n sensors. The set of these sensors is denoted by S . In our proposed model MOUOC, initially, some member CHs (*Member_CH*) are selected randomly. All the nodes have the same probability of becoming the *Member_CH*. These *Member_CH* will participate in the final CH selection process. The CH selection process is based on the local competition. The proposed CH selection algorithm is shown in Algorithm 1. Here, T_h is a predefined threshold used to define the percentage of member CHs in the WSN. The $S_i.C$ and $S_i.R$ represent the chance and competition radius of node i . All the messages are transmitted within distance d_0 . As shown in the algorithm, each *Member_CH* i will prepare a list ($S_i.List$) of nearby *Member_CH* nodes j , such that $d(S_i, S_j) \leq S_i.R$ OR $d(S_i, S_j) \leq S_j.R$. Here, $d(S_i, S_j)$ represents the distance between node i and node j . After preparing the list, each node i will compete with the *Member_CHs* of its list ($S_i.List$) and be selected as the final CH only if it has the highest value of C . Once a node i ($i \in n$) got selected as the final CH, all the nodes in its $S_i.List$ leave the competition for the current round and become a normal node. After receiving the *Final_CH_Msg* messages, all the normal node computes their distance with finally selected CHs. The normal nodes transmit their data directly to the BS if the BS is closer than TN_{MAX} . Otherwise, they will join the nearest CH to form the clusters. The TN_{MAX} is a predefined constant. Therefore, after this phase, we obtain the final CHs and clusters. Our proposed algorithm selects CHs based on local competition, ensuring high scalability in the system. After clusters are formed, data transmission occurs, which is discussed in Phase 2.

Algorithm 1 Algorithm for cluster head selection

```

1: Input: The non-clustered WSN
2: Output: Finally selected Cluster Heads.
3: Set  $S_i.Final\_CH = False \forall i \in n$ 
4:  $\rho \leftarrow rand()$ 
5: if  $\rho \leq T_h$  then
6:    $S_i.Member\_CH = True$ 
7:   Calculate  $S_i.R$  and  $S_i.C$  and share  $Member\_CH\_Msg(ID, Ch, R)$ 
8: end if
9: if  $S_i.Member\_CH = True$  &  $S_j.Member\_CH = True$  then
10:   $S_j$  after getting  $Member\_CH\_Msg$  from  $S_i$ , calculates  $d(S_i, S_j)$ 
11:  if  $d(S_i, S_j) \leq S_j.R$  OR  $d(S_i, S_j) \leq S_i.R$  then
12:    Insert  $S_i$  to  $S_j.List$ 
13:  end if
14: end if
15: if  $S_i.C > S_j.C \forall S_j \in S_i.List$  then
16:   $S_i.Final\_CH = True$  and share  $Final\_CH\_Msg(ID)$ 
17: else
18:   $S_i.Member\_CH = False$  and share  $Quit\_Election\_Msg(ID)$ 
19: end if
20:  $S_j$  after receiving  $Final\_CH\_Msg$  from node  $S_i$ 
21: if  $S_j \in S_i.List$  then
22:   $S_j.Member\_CH = False$  and share  $Quit\_Election\_Msg(ID)$ 
23: end if
24:  $S_j$  after receiving  $Quit\_Election\_Msg(ID)$  from  $S_i$ 
25: if  $S_i \in S_j.List$  then
26:  Remove  $S_i$  from  $S_j.List$ 
27: end if

```

Phase 2: Algorithm for Multi-hop Routing

Our proposed algorithm of data transmission (Algorithm 2) improves data transmission efficiency by leveraging two predefined constants: TN_{MAX} and TH_{MAX} . Normal nodes transmit data directly to the BS if their distance is less than TN_{MAX} ; otherwise, they transmit data to their CH. Similarly,

if a CH is closer than TH_{MAX} to the BS, it transmits data directly to the BS. Otherwise, it will find an optimal intermediate node to transmit the data toward BS. The proposed algorithm selects the CHs closer to the BS and has the minimum cost as an intermediate node. This process is repeated until the data are reached to BS. The details of our proposed algorithm are shown in Algorithm 2.

Algorithm 2 Algorithm for multi-hop routing

```

1: Input: Clustered WSN.
2: Output: Route from CHs to BS is identified.
3: for  $i = 1 : 1 : n$  do
4:   if ( $S_i.Type = N$ ) then
5:     if ( $d(S_i, BS) \leq TN_{MAX}$ ) then
6:       Transmit data to BS directly.
7:     else
8:       Transmit data to  $S_i.CH$ .
9:     end if
10:  end if
11:  if ( $S_i.Type = CH$ ) then
12:    if ( $d(S_i, BS) \leq TH_{MAX}$ ) then
13:      Transmit data to BS directly;
14:    else if  $d(S_i, BS) > TH_{MAX}$  then
15:      Repeat steps until the packet reaches BS
16:      Step 1: Find all nodes  $N_j$  such that
17:         $d(S_i, BS)^2 > d(S_i, N_j)^2 + d(N_j, Sink)^2$ 
18:      Step 2: Calculate effective distances of  $N_j$  as shown below:
19:         $effective\_distance\_N_j = d(S_i, N_j)^2 + d(N_j, Sink)^2$ 
20:      Step 3: Forward packet to  $N_j$  with minimum
21:         $effective\_distance\_N_j$ 
22:    end if
23:  end if

```

We assume that there are n sensors in the WSN. The variable S_i represents the sensor node i ($\forall i \in n$). The variable $S_i.Type$ represents the type of sensor i . A sensor can be of three types. Type N represents the normal node, type CH represents cluster heads and type D is assigned to dead nodes. The variable $S_i.CH$ represents the CH of normal node i . The variable $d(S_i, BS)$ represents the distance between node i and the BS. We manually tuned the value of TH_{MAX} and TN_{MAX} for optimal results and got the values 60 m and 45 m, respectively.

In our evaluation, we have compared the proposed algorithm with four prevalent clustering algorithms, namely, LEACH [1], EEUC [21], FBECS [43], and HROCF [44], across three different scenarios. Our objective is to enhance

the existing clustering algorithms through our proposed model.

By comparing our model with these four algorithms, we aim to assess the effectiveness and performance of our approach in improving clustering efficiency. The comparison results are shown in the section “[Experiment and results](#)”.

Experiment and Results

To simplify our analysis, we assume an ideal MAC layer and transmission links free of errors. Our proposed work, MOUOC, is compared with LEACH [1], EEUC [21], FBECS [43], and HROCF [44] in three scenarios. To

Table 2 Key points of compared algorithms

S. no.	Model name	Parameters used	CH selection	Abbreviation
1	LEACH [1]	None	Probability based	Algo 1
2	EEUC [21]	Distance to BS	Based on RE	Algo 2
3	FBECS [43]	RE, distance to BS, density	Fuzzy logic	Algo 3
4	HROCF [44]	RE, cost and position	Fuzzy logic	Algo 4
5	MOUOC (proposed)	RE, distance to BS, average distance of nearby nodes, standard deviation of the distance of nearby nodes	Fuzzy logic combined with a linear mathematical model	Proposed

Table 3 Parameter

S. no.	Parameter	Value
1	AOI	(0, 0) m to (200, 200) m
2	Sensors (n)	100
3	Location of BS	(100,250)
4	E_{elec}	50 nj
5	ϵ_{fs}	10 pj/bit/m ²
6	ϵ_{mp}	0.0013 pj/bit/m ⁴
7	Packet size	4000 bits
8	EDA	5 nJ/bit/signal
9	d_o	87.7058 m
10	Initial energy (E^0)	0.1 j
11	C (for Algo 2)	0.5
12	T (for Algo 2)	0.4
13	TD_{MAX} (for Algo 2)	90
14	R_{comp}^0 (for Algo 2)	80 m
15	p (for Algo 1)	0.1
16	T_{hr} (for MOUOC)	0.4
17	R_o (for MOUOC)	75 m
18	c_1 (for MOUOC)	0.3
19	c_2 (for MOUOC)	0.2
20	R_{aggr}	10%
21	TH_{MAX}	60
22	TN_{MAX}	45

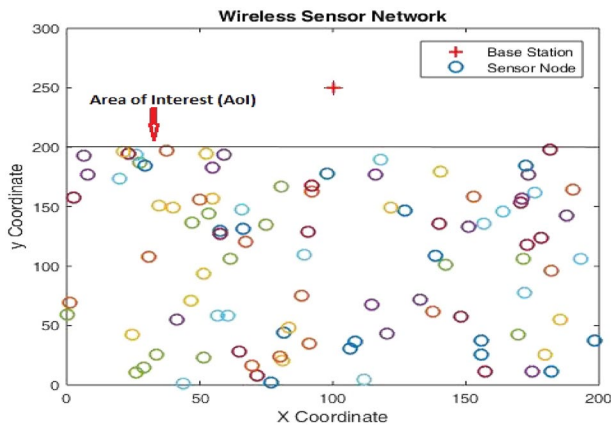


Fig. 9 The WSN for Case 1

conduct a comprehensive assessment, we have specifically chosen these four notable algorithms for comparison. LEACH stands as the pioneering algorithm that introduced the concept of clustering in WSN. EEUC addresses the hot-spot problem by incorporating unequal clustering. On the other hand, FBECs and HROCF utilize fuzzy logic to optimize the clustering process. All the compared algorithms are implemented in MATLAB R2015a. We have used the direct

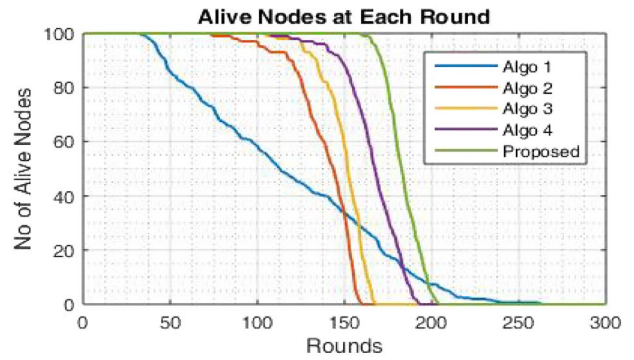


Fig. 10 Remaining nodes versus round for Case 1

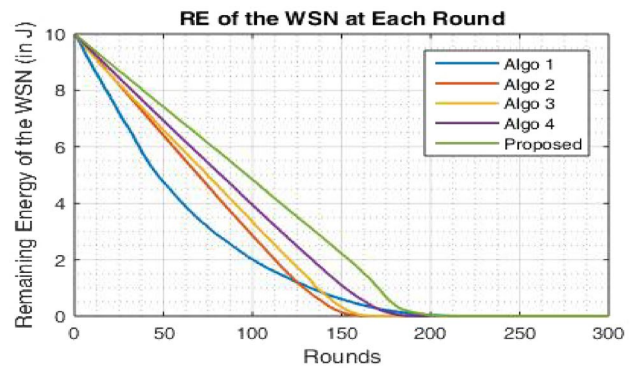


Fig. 11 RE of WSN versus round for Case 1

data transmission method in Algo 1, while multi-hop routing is employed for Algo 2, Algo 3, Algo 4, and the proposed model in all the scenarios. In Algo 3, we have divided the AoI into four equal regions along the y-axis based on their distance to the BS in all the scenarios while keeping other parameters the same as discussed in [43]. By evaluating our proposed algorithm against these established models, we aim to highlight its potential advancements in achieving improved clustering performance. We have used abbreviations for all the models to simplify the result analysis. The details and key aspects of compared models with their abbreviations are shown in Table 2.

To evaluate our proposed model, we have considered three scenarios of WSN. The primary objective of exploring these three scenarios is to comprehensively assess the proposed model’s performance across all possible BS positions. These positions encompass being outside the Area of Interest (AoI), positioned at the corner of the AoI, and located within the AoI itself. The AoI denotes the region where the sensors are deployed. The specific positions of the BS within these scenarios are outlined below

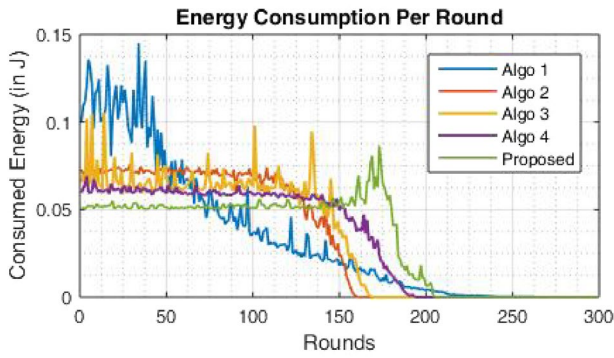


Fig. 12 Energy consumption versus round for Case 1

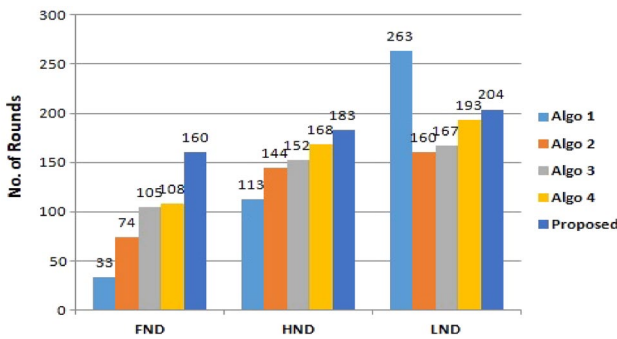


Fig. 13 FND, HND, and LND of all the models for Case 1

- Case 1: Area = 200*200 m², BS = (100, 250), n = 100, E^o = 0.1 J.
- Case 2: Area = 300*300 m², BS = (0, 300), n = 200, E^o = 0.2 J.
- Case 3: Area = 500*500 m², BS = (250, 250), n = 200, E^o = 0.5 J.

Table 3 presents the important parameters utilized in Case 1. For Case 2, and Case 3, the parameters like coverage area, number of nodes, the position of BS and initial energy are changed, and the remaining parameters are the same as in Case 1. The details of the modified parameters are explained at the beginning of each case.

Case 1: Area = 200*200 m², BS = (100,250), n = 100, E^o = 0.1 J.

For Case 1, our coverage area spans 200 × 200 m², extending from location (0, 0) to (200, 200). The BS is situated at (100, 250), outside the AOI, as illustrated in Fig. 9. The plus symbol (+) in the figure denotes the position of BS, the circles (o) represent the position of the sensors, and the lines represent the boundaries of the AOI. The distribution of sensors is random.

Table 4 Sum of RE of nodes across compared algorithms after different rounds for Case 1

Rounds\	Algo 1	Algo 2	Algo 3	Algo 4	Proposed models
50	4.744903	6.402146	6.580201	6.927031	7.410297
100	2.00706	2.854278	3.359092	3.941743	4.835151
150	0.602489	0.133566	0.322308	1.097368	2.221868
200	0.03998	0	0	0	0.023843

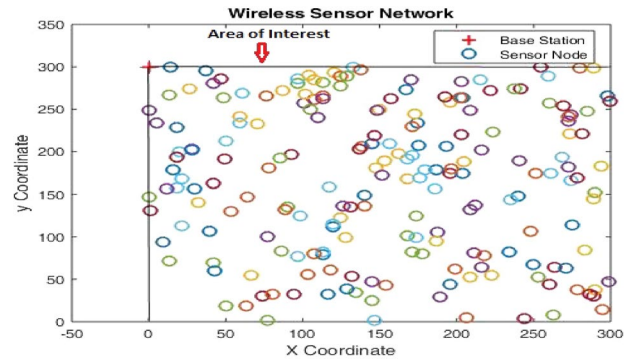


Fig. 14 The WSN for Case 2

To get a complete understanding of the performance of compared models, we have generated several graphs. The comparison of alive nodes in each round of all the models is shown in Fig. 10. It can be observed from the figure that in Algo 1, nodes start dying earliest (from round 33), and after round 200, Algo 1 shows more surviving nodes compared to other algorithms. This is due to the random selection of CHs and direct data transmission, where CHs near the BS consume less energy, allowing them to stay alive, while CHs farther away consume more energy and die earlier. Algo 2 improves Algo 1 by varying the size of clusters for load balancing, so it performs better than Algo 1. The performance of Algo 3 and Algo 4 appears to be similar initially, but after a few rounds, Algo 4 marginally dominates Algo 3. Algo 3 aims to select optimal CHs by considering three parameters: RE, cost, and distance, while Algo 4 improves upon the performance of Algo 3 by considering the communication cost along with the RE and position of each node, so Algo 4 dominates Algo 3. The proposed algorithm performs best by considering the minimum communication cost, adjusting the cluster size based on the node’s position and RE, and maintaining suitable distances between CHs.

Figure 11 depicts the RE of WSN for all the models. The RE of WSN at round *r* is calculated as the total sum of the RE of all alive nodes during round *r*. This comparison is crucial as it provides insights into the rate of energy depletion

in each model, allowing for an assessment of their energy efficiency. It can be observed from the figure that the RE of the WSN decreases significantly in Algo 1 due to its random selection of CHs and direct data transmission. Initially, the performance of Algo 2 and Algo 3 appears to be similar. However, after a few rounds, Algo 3 outperforms Algo 2 due to its superior selection of CHs. The MOUOC exhibits the best overall performance, except for the last few rounds. In these final rounds, alive sensors decrease significantly, leading to inconsistent energy consumption. This trend can also be observed in Fig. 12.

Figure 12 represents the energy consumption of each model in each round, revealing that the energy consumption is highly irregular in Algo 1. In contrast, other models show more consistent energy consumption patterns. It is because Algo 1 selects the CHs randomly, while other models select the CHs based on several parameters. The proposed model consumes the least amount of energy. Still, in the last few rounds, the performance of the proposed algorithm has deteriorated due to the challenge of transmitting data through multi-hop data transmission with only a limited number of remaining nodes.

The comparison of First Node Dead (FND), Half Node Dead (HND), and Last Node Dead (LND) for Case 1 is presented in Fig. 13. The FND, HND, and LND have significant importance in WSNs, because several algorithms, such as [38, 49], consider FND as the lifetime of the WSN, while algorithms like [3, 40] consider HND as the lifetime of the WSN. Although LND is not as significant, we have included it to provide an overall view of the performance of the compared models. Algo 1 performs worst in FND and HND, but it performs best in the case of LND. The LND does not hold much significance. As a significant number of nodes are dead in the sensor network, obtaining accurate information from the remaining sensors is impossible. In most cases, when half of the nodes are dead, the remaining sensors are discarded as well [3]. Figure 13 shows that the proposed algorithm performs 38.2% better

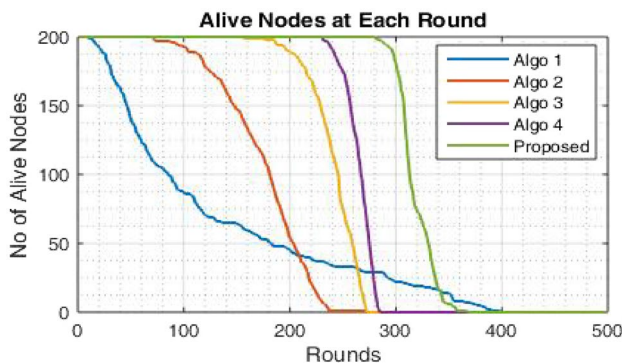


Fig. 15 Remaining nodes versus round for Case 2

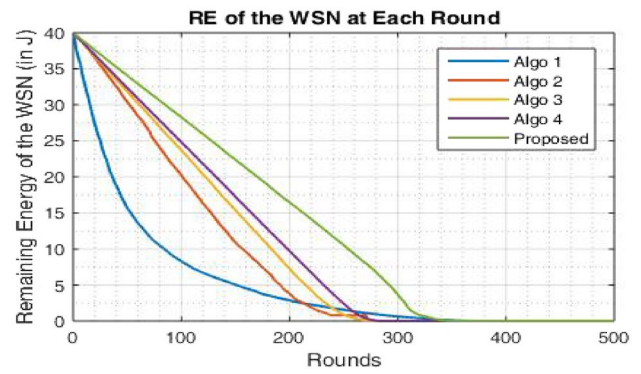


Fig. 16 RE of WSN versus round for Case 2

than Algo 1, 21.3% better than Algo 2, 15.8% better than Algo 3, and 8.2% better than Algo 4 when evaluated using the HND metric. This can be attributed to the selection of optimal CHs, efficient routing and improved load balancing among the nodes.

An analysis is conducted to evaluate the performance of different models by comparing the summation of RE for sensor nodes across multiple rounds, as shown in Table 4. The table presents data for rounds 50, 100, 150, and 200, with values accurately measured up to five decimal places. The analysis reveals that Algo 1 demonstrates the poorest performance during rounds 50, 100, and 150, whereas the proposed model outperforms the others. In round 200, Algo 2, 3, and 4 yield zero values due to the absence of any remaining alive nodes. Algo 1’s comparatively better performance at round 200 is because some nodes near the BS remain alive due to no inter-cluster load. However, this advantage is insignificant, because if more than 50% of nodes are depleted, the WSN is discarded [3].

Case 2: Area = 300*300 m², BS = (000,300), n = 200, E⁰ = 0.2 J.

For this scenario, we have assumed that the BS is positioned at the corner of the AOI. In Case 2, the AoI covers

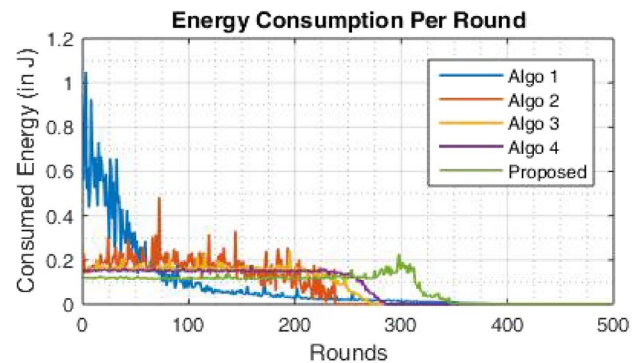


Fig. 17 Energy consumption versus round for Case 2

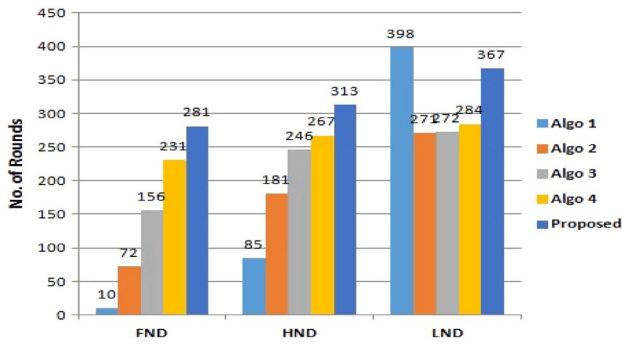


Fig. 18 FND, HND, and LND of all the models for Case 2

Table 5 Sum of RE of nodes across compared algorithms after different rounds for Case 2

Rounds\ models	Algo 1	Algo 2	Algo 3	Algo 4	Proposed
50	15.74612	30.41254	31.87247	32.3678	34.0722
100	8.310691	20.21698	23.70105	24.80552	28.27422
150	5.0345	10.9438	15.41207	17.27071	22.36871
200	2.883021	3.778588	7.285189	9.736677	16.486
250	1.57283	0.797413	0.879234	2.423945	10.55157

an area ranging from (0, 0) to (300, 300) m and comprises 200 sensor nodes distributed randomly. The BS is situated at the corner of the AOI, specifically at coordinates (0, 300), as depicted in Fig. 14. Here, the plus symbol (+) represents the location of the BS, circles (o) represent the location of sensors, and lines show the borders of AoI. The sensors have an initial energy level of 0.2 joules.

In Fig. 15, the comparison of alive nodes in each round of each model is shown. The sensors in Algo 1 start depleting at the earliest. This is attributed to the random selection of CHs and the direct data transmission method used in Algo 1. Algo 2 outperforms Algo 1 due to its improved load balancing achieved by varying the size of clusters. The performance of Algo 3 and Algo 4 is consistent; still, Algo 4 outperforms Algo 3 due to better CHs selection and load balancing. However, the proposed algorithm surpasses all others due to optimal CHs selection, improved load balancing, and efficient data routing.

Figure 16 provides a comparative analysis of the RE in WSNs across different models. This comparison is crucial as it sheds light on the rate of energy depletion in each model, enabling an assessment of their energy efficiency. Notably, the figure highlights the distinctive behavior of Algo 1. Initially, Algo 1 experiences a significant drop in the energy of WSN, but it stabilizes when only a few nodes remain alive. This phenomenon is attributed to the direct transmission employed in Algo 1, which leads to higher energy

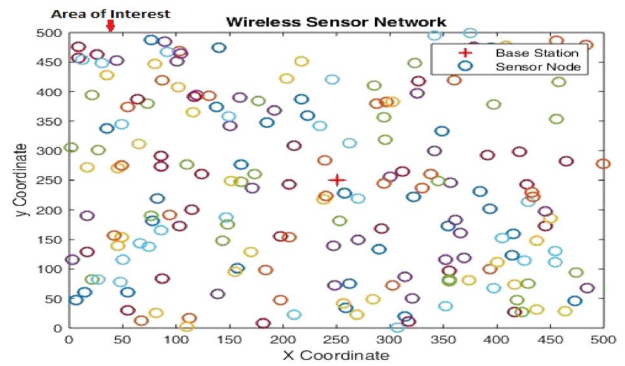


Fig. 19 The WSN for Case 3

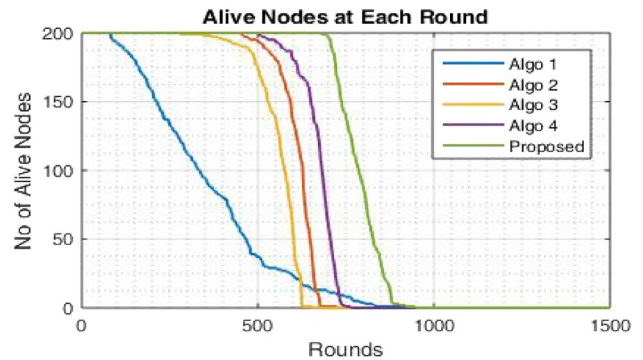


Fig. 20 Remaining nodes versus round for Case 3

consumption by CHs located farther from the BS, causing them to exhaust their energy sooner. However, energy consumption decreases as most sensors far from the BS become dead, leaving only the sensors closer to the BS alive. The proposed algorithm performs the best due to the optimal selection of CHs, improved load balancing, and efficient data routing.

The diagram presented in Fig. 17 illustrates the energy consumption of each model in every round. Upon examination, it becomes evident that Algo 1 exhibits high inconsistency in energy consumption, whereas Algo 2, Algo 3, Algo 4, and the proposed algorithm demonstrate more consistent energy consumption patterns. Particularly, the proposed algorithm showcases the lowest energy consumption in each round, except for the last few rounds. This achievement can be attributed to the proposed algorithms' emphasis on improved CH selection, balanced load distribution among CHs and normal nodes, as well as maintaining optimal distances between CHs. As a result, it becomes the most energy-efficient model, with the exception of the last few rounds. However, in the last few rounds, the energy consumption of the proposed algorithm is relatively higher compared to other algorithms due to the fact that in those rounds,

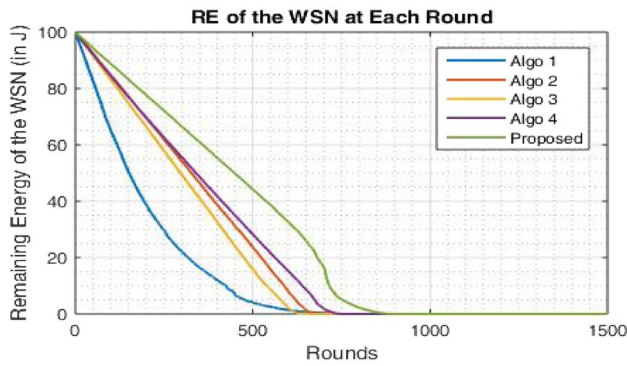


Fig. 21 RE of WSN versus round for Case 3

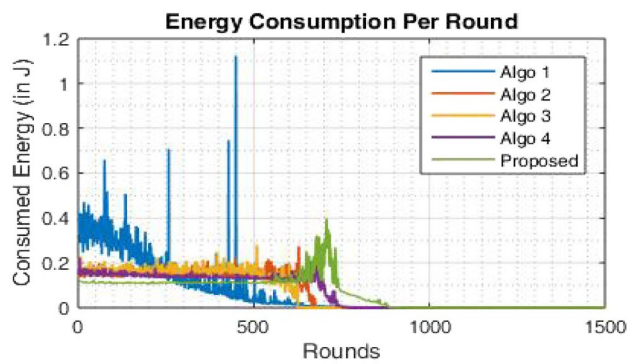


Fig. 22 Energy consumption versus round for Case 3

a majority of the sensors in other algorithms have already depleted their energy. In contrast, a significant number of sensors are still active in the proposed algorithm. Consequently, the active nodes continue to consume energy, causing the energy consumption of the proposed model to be the highest during this phase.

Figure 18 compares the FND, HND, and LND of all algorithms in Case 2. The figure clearly shows that the proposed model outperforms all other algorithms in terms of FND and HND. Specifically, the proposed algorithm performs 72.8% better than Algo 1, 42.2% better than Algo 2, 21% better than Algo 3, and 14.7% better than Algo 4 when evaluated using the HND metric. On the other hand, Algo 1 exhibits better performance in terms of LND, because some nodes near the BS remain alive due to no inter-cluster load. These findings emphasize the potential of load balancing and efficient routing in improving the lifetime of WSNs, especially in terms of FND and HND.

We conducted an analysis of the summation of RE for alive sensor nodes across multiple rounds in all models, as depicted in Table 5. The table provides comparative data for rounds 50, 100, 150, 200, and 250, with values accurate up to five decimal places. The results demonstrate that

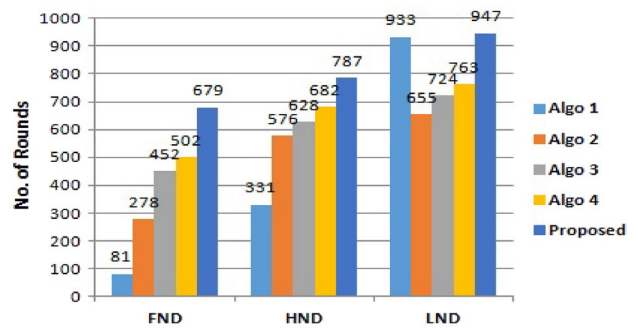


Fig. 23 FND, HND, and LND of all the models for Case 3

Table 6 Sum of RE of nodes across compared algorithms after different rounds for Case 3

Rounds\ models	Algo 1	Algo 2	Algo 3	Algo 4	Proposed
100	65.09083	83.16402	84.77899	84.4875	88.70867
200	38.71584	66.36428	69.44261	69.69816	77.65009
300	22.11478	49.41275	54.19218	55.63196	66.53523
400	11.95578	32.64867	38.73252	41.77027	55.40596
500	4.136032	15.91745	23.79679	28.28215	44.26891

Algo 1 exhibits the poorest performance, while the proposed model outperforms the others. Furthermore, as the number of rounds increases, the performance gap between existing and proposed models widens. These findings strongly support the notion that the optimal selection of CHs, effective load balancing, and optimal distance between CHs can significantly enhance energy consumption.

Case 3: Area = 500*500 m², BS = (250,250), n = 200, E⁰ = 0.5 J.

In this scenario, we have carefully chosen an area measuring 500 × 500 m² and positioned the BS at the coordinates (250, 250). Figure 19 illustrates a snapshot of the AOI and the position of BS in Case 3. In this figure, the circles (o) represent the position of sensors, and the plus symbol (+) represents the position of BS. To evaluate the models' performance, we have compared them on several criteria. Figure 20 represents the number of alive nodes in each round of each model. The figure shows that Algo 1 performs the worst in terms of FND, while the proposed MOUOC algorithm performs the best. This occurs, because, in Algo 1, the selection of CHs is solely based on probability, which occasionally results in suboptimal CHs' selection and, consequently, high energy consumption by the sensors. Additionally, Algo 1 employs direct data transmission from CHs to the BS, resulting in higher energy consumption and faster depletion for CHs located far from the BS. Algo 2 surpasses Algo 1 by implementing

improved load balancing through varying cluster sizes. Algo 4 outperforms Algo 3 due to superior CH selection and load balancing. In our proposed model, we have endeavored to further enhance these algorithms by selecting better CHs and improving load balancing efficiently. As a result, our model consumes the least energy, enabling sensors to remain active for a longer duration.

Figure 21 compares the RE of the WSN. The RE of the WSN at round r is calculated as the sum of the RE of all the active nodes at that round. From the figure, it can be observed that the RE of the WSN decreases significantly in Algo 1. Due to the random selection of CHs and the utilization of direct data communication in Algo 1, the CHs located far from the BS experience higher energy consumption and deplete earlier. Initially, the difference in RE between Algo 2, Algo 3, and Algo 4 is not substantial, but after a few rounds, the difference becomes significant, with Algo 4 outperforming Algo 1, Algo 2, and Algo 3. The proposed model exhibits the best performance due to its optimized energy consumption and load-balancing methods.

To gain an overview of each model's performance throughout its lifetime, we compared the energy consumption of each model in every round, as shown in Fig. 22. Upon observing the figure, it becomes apparent that energy consumption in Algo 1 is highly inconsistent. The irregular lines in the graph for Algo 1 indicate that, in certain rounds, energy consumption is considerably higher, surpassing the energy consumption of previous rounds. This inconsistency arises due to the fact that, in Algo 1, CHs are selected solely based on probability, and direct data transmission proves to be ineffective in larger areas. In contrast, Algo 2, Algo 3, Algo 4, and the proposed model demonstrate consistent energy consumption throughout their lifetimes. Notably, the proposed model excels in achieving the highest level of consistency and consuming the least amount of energy. This can be attributed to several factors, including the optimal selection of CHs, optimal cluster size, efficient load balancing among sensors, and effective data routing from CHs to the BS.

The comparison of FND, HND, and LND among all the models is presented in Fig. 23. The figure shows that Algo 1 performs the poorest in terms of FND and HND, whereas the proposed algorithm demonstrates the best performance. In this scenario, the proposed model performs 57.9% better than Algo 1, 26.8% better than Algo 2, 20.2% better than Algo 3, and 13.3% better than Algo 4 when evaluated using the HND metric. These results support our assumption that selecting optimal CHs, implementing better load-balancing techniques, and optimizing routing mechanisms can enhance the network's lifetime.

The summation of the RE for the alive nodes has been examined across various rounds for all models, as depicted

in Table 6. In this table, we present the comparative data for rounds 100, 200, 300, 400, and 500. The values provided are accurate up to five decimal places. The results indicate that Algo 1 exhibits the poorest performance, while the proposed model demonstrates the best performance. For instance, at round 100, the MOUOC outperforms Algo 1 by 23.6%, Algo 2 by 6%, Algo 3 by 4.4%, and Algo 4 by 4.7%. This performance gap widens with an increasing number of rounds. These findings affirm that the optimal selection of CHs, efficient load balancing, optimal distance between CHs, and efficient routing significantly enhance energy consumption.

Based on the analysis of these scenarios, it is observed that the performance of the proposed model varies with the position of the BS and the distribution of the sensors. However, despite these variations, the overall findings confirm that the proposed model's approach, which focuses on optimal CHs' selection, load balancing, maintaining an optimal distance between CHs and efficient routing, plays a crucial role in extending the lifespan of the network.

Conclusion and Future Work

This paper aims to introduce a new and innovative model that optimizes the selection of CHs, balances the load, and enables efficient multi-hop data transmission in WSNs. Our proposed model operates in a distributed manner and is designed to be lightweight enough for deployment on actual sensors. To achieve these objectives, we have developed a CH selection algorithm that leverages both lightweight fuzzy logic and a linear mathematical model. By incorporating the latter, we are able to simplify the fuzzy logic, improving its efficiency without sacrificing performance.

We evaluated our proposed model against four alternative models in three distinct scenarios, using several metrics, including total alive nodes, the RE of the WSN, energy consumption, and the round of FND, HND, and LND occurrences. The results confirm that our algorithm effectively balances energy consumption, improves network lifetime, and addresses the hot-spot issue using optimal competition radii for CHs.

The parameters $c1$ and $c2$, which are used to calculate the competition radius and value of the maximum competition radius, are currently manually tuned for improved results. However, in the future, some advanced machine learning-based methods may be developed to determine the optimal values of these parameters based on specific circumstances. Moreover, while we have endeavored to use the best parameters for CH selection, there is always room for improvement. Finally, our proposed model currently considers only one stationary BS. However, in the future, it may be extended to support more than BS or a mobile BS.

Funding Not applicable.

Declarations

Conflict of Interest Not applicable.

Ethics Approval The authors of this article did not conduct any studies involving human or animal subjects.

Financial Interests The authors do not have any pertinent financial or non-financial disclosures to make.

References

- Heinzelman WR, Chandrakasan A, Balakrishnan H. Energy-efficient communication protocol for wireless microsensor networks. In: Proceedings of the 33rd Annual Hawaii International Conference on System Sciences. IEEE; 2000. p. 10.
- El Alami H, Najid A. Fuzzy logic based clustering algorithm for wireless sensor networks. In: Sensor technology: concepts, methodologies, tools, and applications. IGI Global; 2020. p. 351–71
- Sert SA, Bagci H, Yazici A. Mofca: multi-objective fuzzy clustering algorithm for wireless sensor networks. *Appl Soft Comput*. 2015;30:151–65.
- Zhou Y, Fang Y, Zhang Y. Securing wireless sensor networks: a survey. *IEEE Commun Surv Tutor*. 2008;10(3):6–28.
- Want R, Farkas KI, Narayanaswami C. Guest editors' introduction: energy harvesting and conservation. *IEEE Pervasive Comput*. 2005;4(1):14–7.
- Roy NR, Chandra P. Energy dissipation model for wireless sensor networks: a survey. *Int J Inf Technol*. 2020;12:1343–53.
- Ould Amara S, Beghdad R, Oussalah M. Securing wireless sensor networks: a survey. *EDPACS*. 2013;47(2):6–29.
- Estrin D, Govindan R, Heidemann J, Kumar S. Next century challenges: Scalable coordination in sensor networks. In: Proc. Int. Conf. Mobile Computing and Networking (MOBICOM). 1999. p. 263–70.
- Sheng Z, Yang S, Yu Y, Vasilakos AV, McCann JA, Leung KK. A survey on the IETF protocol suite for the internet of things: standards, challenges, and opportunities. *IEEE Wirel Commun*. 2013;20(6):91–8.
- Jose J, Kumar SM, Jose J. Energy efficient recoverable concealed data aggregation in wireless sensor networks. In: 2013 IEEE international conference on emerging trends in computing, communication and nanotechnology (ICECCN). IEEE; 2013. p. 322–29.
- Shi S, Liu X, Gu X. An energy-efficiency optimized leach-c for wireless sensor networks. In: 7th international conference on communications and networking in China. IEEE; 2012. p. 487–92.
- Liao Q, Zhu H. An energy balanced clustering algorithm based on leach protocol. *Appl Mech Mater*. 2013;341:1138–43.
- Kurumbanshi S, Rathkanthiwar S. Increasing the lifespan of wireless Adhoc network using probabilistic approaches: a survey. *Int J Inf Technol*. 2018;10:537–42.
- Ye M, Li C, Chen G, Wu J. Eecs: an energy efficient clustering scheme in wireless sensor networks. In: PCCC 2005. 24th IEEE international performance, computing, and communications conference. IEEE; 2005. p. 535–40.
- Soro S, Heinzelman WB. Prolonging the lifetime of wireless sensor networks via unequal clustering. In: 19th IEEE international parallel and distributed processing symposium. IEEE; 2005. p. 8.
- Yadav M, Bhola A, Jha CK. A novel WSN protocol for increasing network life using combination of node's positions and communication range. *Int J Inf Technol*. 2020;12(1):77–84.
- Sharma R, Singh U. Fuzzy based energy efficient clustering for designing WSN-based smart parking systems. *Int J Inf Technol*. 2021;13:2381–7.
- Devika G, Ramesh D, Karegowda AG. Energy optimized hybrid PSO and wolf search based leach. *Int J Inf Technol*. 2021;13:721–32.
- Ramesh D, Karegowda AG. Firefly and grey wolf search based multi-criteria routing and aggregation towards a generic framework for leach. *Int J Inf Technol*. 2022;14(1):105–14.
- Sinha A, Lobiyal DK. Performance evaluation of data aggregation for cluster-based wireless sensor network. *Hum Centric Comput Inf Sci*. 2013;3:1–17.
- Li C, Ye M, Chen G, Wu J. An energy-efficient unequal clustering mechanism for wireless sensor networks. In: IEEE international conference on mobile Adhoc and sensor systems conference. IEEE; 2005. p. 8.
- Shurman M, Awad N, Al-Mistarihi MF, Darabkh KA. Leach enhancements for wireless sensor networks based on energy model. In: 2014 IEEE 11th international multi-conference on systems, signals and devices (SSD14). IEEE; 2014. p. 1–4.
- Younis O, Fahmy S. Heed: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. *IEEE Trans Mob Comput*. 2004;3(4):366–79.
- Elshrkawey M, Elsherif SM, Wahed ME. An enhancement approach for reducing the energy consumption in wireless sensor networks. *J King Saud Univ Comput Inf Sci*. 2018;30(2):259–67.
- Pandey SK, Patel S, et al. Energy distance clustering algorithm (EDCA) for wireless sensor networks. In: 2019 9th international conference on cloud computing, data science and engineering (confluence). IEEE; 2019. p. 287–92.
- Singh B, Lobiyal DK. A novel energy-aware cluster head selection based on particle swarm optimization for wireless sensor networks. *Hum Centric Comput Inf Sci*. 2012;2(1):1–18.
- Coello CC, Lechuga MS. Mopso: a proposal for multiple objective particle swarm optimization. In: Proceedings of the 2002 congress on evolutionary computation. CEC'02 (Cat. No. 02TH8600). IEEE; 2002. p. 1051–56.
- Yang E, Erdogan AT, Arslan T, Barton N. An improved particle swarm optimization algorithm for power-efficient wireless sensor networks. In: 2007 ECSIS symposium on bio-inspired, learning, and intelligent systems for security (BLISS 2007). IEEE; 2007. p. 76–82.
- Wu X, Lei S, Jin W, Cho J, Lee S. Energy-efficient deployment of mobile sensor networks by pso. In: Advanced web and network technologies, and applications: APWeb 2006 international workshops: XRA, IWSN, MEGA, and ICSE, Harbin, China, January 16–18, 2006. Proceedings 8. Springer; 2006. p. 373–82.
- Sahoo BM, Amgoth T. An improved bat algorithm for unequal clustering in heterogeneous wireless sensor networks. *SN Comput Sci*. 2021;2(4):290.
- Rambabu C, Prasad V, Prasad KS. A new version of energy-efficient optimization protocol using ICMA-PSOGA algorithm in wireless sensor network. *SN Comput Sci*. 2022;3(5):353.
- Manasa P, Shaila K, Venugopal KR. Energy concerned clustering mechanism to ensure reliable data transmission in wireless sensor network. *SN Comput Sci*. 2021;2(4):266.
- Lekhray, Kumar A, Kumar A. Multi criteria decision making based energy efficient clustered solution for wireless sensor networks. *Int J Inf Technol*. 2022;14(7):3333–42.
- Rajpoot P, Dwivedi P. Optimized and load balanced clustering for wireless sensor networks to increase the lifetime of WSN using MADM approaches. *Wirel Netw*. 2020;26(1):215–51.

35. Choudhary A, Nizamuddin M, Sachan VK. A hybrid fuzzy-genetic algorithm for performance optimization of cyber physical wireless body area networks. *Int J Fuzzy Syst.* 2020;22:548–69.
36. Sheleba A, Tabatabaei S. A novel method for clustering in WSNS via topsis multi-criteria decision-making algorithm. *Wirel Pers Commun.* 2020;112(2):985–1001.
37. Pandey SK, Singh B. Topsis-based optimal cluster head selection for wireless sensor network. *Res Rep Comput Sci.* 2023;2(3):77–86.
38. Kim J-M, Park S-H, Han Y-J, Chung T-M. Chef: cluster head election mechanism using fuzzy logic in wireless sensor networks. In: 2008 10th international conference on advanced communication technology, 1. IEEE; 2008. p. 654–59.
39. Sood P. A fuzzy logic based clustering algorithm for WSN to extend the network lifetime. *J Glob Res Comput Sci.* 2018;9(6):19–25.
40. Bagci H, Yazici A. An energy aware fuzzy approach to unequal clustering in wireless sensor networks. *Appl Soft Comput.* 2013;13(4):1741–9.
41. Khan BM, Bilal R, Young R. Fuzzy-topsis based cluster head selection in mobile wireless sensor networks. *J Electr Syst Inf Technol.* 2018;5(3):928–43.
42. Logambigai R, Kannan A. Fuzzy logic based unequal clustering for wireless sensor networks. *Wirel Netw.* 2016;22:945–57.
43. Mehra PS, Doja MN, Alam B. Fuzzy based enhanced cluster head selection (FBECS) for WSN. *J King Saud Univ Sci.* 2020;32(1):390–401.
44. Kumaran RS, Bagwari A, Nagarajan G, Kushwah SS. Hierarchical routing with optimal clustering using fuzzy approach for network lifetime enhancement in wireless sensor networks. *Mob Inf Syst.* 2022;2022:11. <https://doi.org/10.1155/2022/6884418>.
45. Heinzelman WB, Chandrakasan AP, Balakrishnan H. An application-specific protocol architecture for wireless microsensor networks. *IEEE Trans Wirel Commun.* 2002;1(4):660–70.
46. Manjeshwar A, Agrawal DP. Teen: A routing protocol for enhanced efficiency in wireless sensor networks. In: *Ipdps 1.* 2001; 189.
47. Lindsey S, Raghavendra CS. Pegasus: power-efficient gathering in sensor information systems. In: *Proceedings, IEEE aerospace conference, 3.* IEEE; 2002. p. 3.
48. Mamdani EH. Application of fuzzy logic to approximate reasoning using linguistic synthesis. *IEEE Trans Comput.* 1977;100(12):1182–91. <https://doi.org/10.1109/TC.1977.1674779>.
49. Karl H, Willig A. *Protocols and architectures for wireless sensor networks.* New York: Wiley; 2007.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.