IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 30 December 2023, accepted 14 January 2024, date of publication 13 February 2024, date of current version 21 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3365490

RESEARCH ARTICLE

Improvement of Wireless Sensor Network Lifetime via Intelligent Clustering Under Uncertainty

LAXMINARAYAN SAHOO[1], SHRI SUPRIYAN SEN[1], KALISHANKAR TIWARY[2], SARBAST MOSLEM[3], AND TAPAN SENAPATI[4]

[1]Department of Computer and Information Science, Raiganj University, Raiganj 733134, India
[2]Department of Mathematics, Raiganj University, Raiganj 733134, India
[3]School of Architecture, Planning and Environmental Policy, University College Dublin, Belfield, Dublin, D04 V1W8 Ireland
[4]School of Mathematics and Statistics, Southwest University, Beibei, Chongqing 400715, China

Corresponding authors: Sarbast Moslem (sarbast.moslem@ucd.ie) and Tapan Senapati (math.tapan@gmail.com)

This article was supported in part by the Science Foundation Ireland (SFI) through the Digital Voting Hub for Sustainable Urban Transport System, VOTE-TRA Project, under Grant 22/NCF/DR/11309.

ABSTRACT This research proposes a novel framework that integrates intelligent clustering algorithms with "multi-criteria decision-making (MCDM)" techniques to enhance the longevity of WSNs in uncertain environment. Clustering techniques are crucial in WSNs for data aggregation and energy-efficient communication. To create energy efficient network, the proposed framework incorporates intelligent clustering algorithms that perform clustering dynamically in the presence of uncertain parameter and employed MCDM techniques to select of energy efficient CHs for clustering. The intelligent clustering algorithms employ data-driven approaches, machine learning and optimization algorithms to create optimal cluster formation, cluster head selection and energy efficiency. An intelligent clustering mechanism has been made using the Silhouette Index (SI) score. Utilizing the SI score as a benchmark, we conducted optimized clustering with the "Density-Based Spatial Clustering of Applications with Noise (DBSCAN)" algorithm. We employed the elbow method to validate the SI score in conjunction with the k-Means clustering algorithm. By considering uncertainty factors in the decision-making process, the proposed algorithms can effectively adapt the network's operation to changing conditions, thus improving the overall lifetime of the WSN. Furthermore, the framework integrates MCDM approaches to prioritize cluster formation and cluster head selection criteria. Triangular Fuzzy Numbers are compatible with fuzzy logic systems, which are designed to handle uncertainty and imprecision. The triangular shape aligns well with the concept of fuzzy sets and fuzzy reasoning. Due to this reason TFNs have been considered to represent uncertain parameters. In the end, an experiment relating to WSNs has been studied and the results have been visually presented. It has been noticed that the suggested approach outperformed the "residual energy-aware clustering with isolated nodes (REAC-IN)" model, "Low-Energy Adaptive Clustering Hierarchy Fuzzy Clustering (LEACH-FC)" and "hybrid energy efficient distributed (HEED)" by 38%, 15% and 43%, respectively. The PSO and BFAO applied optimized clustering has been outperformed by 35% and 22%, respectively. To verify the simulation results, testing of hypotheses has been conducted.

INDEX TERMS Wireless sensor networks, lifetime extension, cluster head selection, residual energy, DBSCAN, MCD.

I. INTRODUCTION

Much interest has grown in wireless sensor networks (WSNs) recently. This rising curiosity has required in-depth information, giving researchers a strict comprehension of this

The associate editor coordinating the review of this manuscript and approving it for publication was Francisco J. Garcia-Penalvo.

study area. [1]. A WSN is a network comprising a few or more sensor units that cooperate to complete particular tasks, like sensing the physical environment, forming judgments, and sending the sensed data to the proper destination (see Fig. 1) [2], [3]. Due to its use in various sectors, such as agriculture, the military defence sector, home appliances, remote sensing, etc., wireless communication has several advantages. The WSNs are used for communication in several security-based applications, including those that monitor the environment, vehicle traffic, smart offices, and battlefield surveillance [4], [5]. There aren't many WSN restrictions, such as memory units, limited-energy modules, and high-performance times. Rather [6], it has been the subject of ongoing research for a few applications [7]. Each sensor node within a WSN consists of four essential components: transceivers, sensors, power sources, and microcontrollers. These sensors continuously record the required parameters, and the microcontroller processes the collected data, transmitting it to the ground station via the communication unit, either directly through a single node or through adjacent nodes [8]. Due to the inherent limitations and challenges in recharging node power sources, research has increased to enhance energy balance and efficiency in WSNs [9]. Sensor devices have a finite lifespan, prompting efforts to extend their operational longevity by developing energy-efficient routing protocols [10], [11].

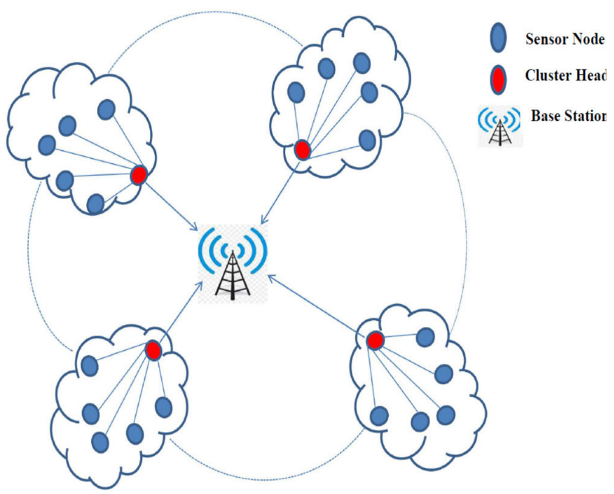

FIGURE 1. Overall framework of WSNs.

Routing is the distinguishing hallmark of WSNs compared to other ad hoc wireless networks. When it comes to WSNs, energy-efficient routing algorithms are necessary for relaying observed data from "sensor nodes (SNs)" to the "base station (BS)". This is necessary to ensure the network will remain operational for an extended period. Most WSN variants rely on cluster-based protocols to minimize energy consumption in sensor nodes. These clustering algorithms within WSNs leverage various optimization techniques to manage clustering operations efficiently. Optimization methods such as "Artificial Bee Colony (ABC)" [12], "Particle Swarm Optimization (PSO)" [13], and the "Bacterial Foraging Algorithm for Optimization (BFAO)" [14], among others, have been employed to optimize cluster numbers. However, when compared to "Low Energy Adaptive Clustering Hierarchy (LEACH)" and other optimization techniques, the adoption of these optimization algorithms has yielded reduced energy consumption, improved network efficiency, and enhanced data delivery performance. In uncertain environment of WSNs the parameters are uncertain and possess fuzzy values, indicating imprecision and variability, often requiring fuzzy logic or other uncertainty modeling approaches for accurate representation. Incorporating uncertainty considerations in WSNs is essential for designing robust, adaptive, and reliable systems capable of operating effectively in dynamic and unpredictable environments. Strategies for handling uncertainty include the use of fuzzy logic, probabilistic models, machine learning algorithms, and adaptive protocols to ensure the resilience and adaptability of WSNs. In this research, uncertainty is represented using triangular fuzzy numbers, and the clustering process is executed through the utilization of intelligent clustering, which relies on the DBSCAN-based machine learning algorithm. In many real-world situations, uncertainty is often characterized by a range of possible values with a most likely or typical value in the middle. TFNs naturally capture this type of uncertainty, making them a suitable choice for modeling uncertainties in various domains. Triangular fuzzy numbers require only three parameters, which mean they may be easier to elicit or estimate when compared to other types of fuzzy numbers that might need more data points. This can be particularly advantageous when dealing with limited or imprecise data. The simplicity and interpretability of triangular fuzzy numbers make them suitable for decision-making processes where uncertainties need to be considered. Decision-makers can use triangular fuzzy numbers to represent uncertain information and make decisions based on a more comprehensive understanding of the uncertainty involved. TFNs are more suitable for representing discrete or interval uncertainty. For continuous uncertainty, other fuzzy number representations like trapezoidal [15], Gaussian fuzzy numbers [16] or generalized bell-shaped fuzzy number [17] may be more appropriate. In this study the uncertainty is discrete and it is represented by range of possible values most likely middle value. Based on these requirements the uncertainty has been represented by TFNs. In this study, triangular fuzzy numbers (TFNs) have been used to signify ambiguous parameters due to its mentioned advantages.

MCDM techniques are essential for the selection of CHs in clustering within WSNs due to their ability to handle multiple criteria, address conflicting objectives, integrate subjective, objective criteria, provide flexibility and optimize network performance in dynamic and uncertain environments. These techniques contribute to the efficiency, adaptability and sustainability of WSNs ensuring effective CH selection for enhanced overall network performance and lifetime. This study incorporates various parameters for the selection of

CHs. Leveraging the capability of MCDM techniques to handle multiple criteria's, CHs are chosen from several cluster nodes based on the mentioned properties.

In this paper, the uncertain parameter is represented by TFNs and optimized number of clusters has been obtained by SI and verified by the elbow method. Clustering is carried out by the DBSCAN based intelligent clustering technique for the first iteration. For the successive round/iteration, we have developed an algorithm which will choose the cluster heads (CHs) based on several WSNs parameters/characteristics (see Fig. 2). Acronyms used in this paper have been provided in Table 1.

TABLE 1. List of abbreviations.

Abbreviations	Explanation
ABC	Artificial Bee Colony
BFAO	Bacterial foraging algorithm for optimization
BS	Base station
CHs	Cluster heads
DBSCAN	Density-based spatial clustering of applications with noise
HEED	"Hybrid energy efficient distributed"
LEACH	"Low energy adaptive clustering hierarchy"
LEACH-FC	"Low-energy adaptive clustering hierarchy fuzzy clustering"
MCDM	"Multi-criteria decision-making"
PSO	"Particle swarm optimization"
REAC-IN	"Residual energy-aware clustering with isolated nodes"
SI	Silhouette index
SN	Sensor node
TFNs	Triangular fuzzy numbers
WSNs	Wireless sensor networks
WCSS	Within Cluster Sum of Squares

A. RELATED WORKS

In 2008, Leu et al. [18] developed an energy-conscious clustering approach that considered the presence of isolated nodes during the process. The selection of CHs in the WSN was determined by the "Regional Energy-Aware Clustering with Isolated Nodes (REAC-IN)" model, assigning weight to each node. The weight parameter was calculated based on factors such as the average regional energy of sensors within the group and the residual energy of individual nodes. Improper clustering algorithm implementation shows node isolation and inability to reach the CHs problem in WSNs. The separated nodes then communicate via sink node instead of CHs, which consumes more residual energy, leading to a network lifetime problem. A node has been identified as an isolation node to extend the network lifetime by considering several features like the average energy in the area and the distance between sensor nodes and sink nodes. The REAC-IN results were examined differently and contrasted with the currently used clustering techniques. It has been found that REAC-IN outperformed

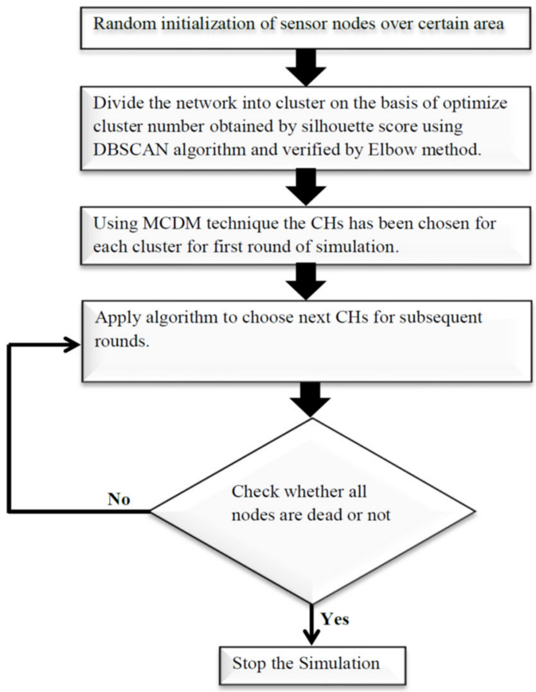

FIGURE 2. Flowchart depicting the suggested approach employed in this study.

some of the very well-known clustering algorithms. In 2013, Solaiman et al. [19] examined the effectiveness of PSO-based energy optimization for WSNs and highlighted the manuscript's observations from a theoretical perspective. However, contrasting those above, in 2015, the study conducted by Parvin et al. [20] demonstrated a comparable pattern of quantitative investigation coupled with a distinct methodological approach. The conceptualization of the designated design modeling is founded on PSO clustering. During this phase, PSO clustering examines the optimal route establishment utilizing the search technique. The devised system exhibits a certain level of superiority when subjected to simulation with the NS2 simulator. In addition to PSO-based optimization, many studies have evaluated the efficacy of "bacterial foraging optimization (BFO)" in data delivery and energy consumption for clustering in WSN. In 2015, Moharamkhani et al. [21] developed an innovative BFO-based approach to improve the consumption of energy by WSNs. The author states that BFO might be used to solve additional multidimensional problems and cluster data energy-efficiently. The comparative performance research revealed that the proposed system outperforms LEACH regarding energy performance. In 2018, Murugan and Sarkar [22] developed the "firefly cyclic grey wolf optimization (FCGWO)". Compared to LEACH and PSO, the BFAO's design approach achieved higher energy performance and Quality of Service (QoS). The study also looked at bottleneck conditions caused by the design constraints of PSO- and BFAO-based clustering problem formulations. In the year 2019, the work of Sambo et al. [23] brought forth advanced methods in machine learning and computational intelligence. They organized computational techniques into distinct categories, including swarm intelligence, fuzzy logic, neural networks, genetic algorithms, and reinforcement learning, to reflect how computational intelligence is applied. Uses of these computational intelligence applications, like scalability, data transmission speed, and data aggregation, have been examined. Furthermore, we observed that these methods extended the network's longevity and enhanced service quality. The integration of the hybrid model also decreased network interference. In 2020, Lata et al. [24] examined the application of the LEACH method and the fuzzy principle for clustering. A new variation of LEACH, LEACH-FC, has been developed for the CH selection. The proposed strategy employed a centralized method because fuzzy logic was used to choose the appropriate CH for clustering. Because the method's cluster formation and CH selection are not distributed, load balancing can be carried out successfully. The network lifetime and the node's energy were balanced with the other parameters. In 2020, Xiuwu et al. [25] created a clustering framework called CLWPA for the heterogeneous WSN. The sensor nodes were clustered using the CLWPA method. To deploy the best nodes in the network, they initially modified the wolf pack algorithm (WPA). The proposed CLWPA model was then developed by combining the DEEC and heterogeneous network routing algorithms. Three currently used routing protocols were considered for the comparison, and the simulation was run under various scenarios. The WPA enhanced local search functionality and kept optimization from settling on local solutions. In 2020, Deepa et al. proposed a novel PSO [26] and BFAO clustering algorithm [14]. The research issue within the framework of energy-aware clustering is investigated from many angles. Although LEACH-based clustering policies provide some energy-aware data aggregation in one-hop or multi-hop hierarchical solutions, there is a problem with data delivery reliability. As a result, from an energy perspective, it eventually exhausts crucial WSN computing resources. Several effective optimization-based clustering strategies have been used to overcome this barrier; however, the present PSO and BFAO solutions require a substantial amount of computationally intensive recursive and iterative procedures. As a result, it is challenging to find a perfect solution with faster process execution when using PSO and BFAO-based clustering algorithms. Due to its larger particle computation, PSO also increases computational complexity. Furthermore, regarding the trade-off between data transmission performance and WSN energy efficiency, PSO and BFAO do not impose flawless solutions. As a result, the WSN energy and clustering problem is still a problem. Here, by considering criteria like the distance from the sink, the average distance between cluster nodes, the cluster's reliability, and residual energy, we have employed a novel technique to increase the lifetime of WSNs. It has been known that poor or risky communication can have highly detrimental impacts, including an increase in communication noise and a negative impact on the sensor nodes' batteries. Batteries and residual energy are closely related, which will reduce the network's lifespan and prolong network delay.

The findings of the experiment have been investigated and represented at the end of the study. The REAC-IN model, LEACH-FC and HEED models have been outperformed by the proposed technique by 38%, 15%, and 43%, respectively. Clustering performance has been outperformed by 35% and 22%, respectively, compared to PSO and BFAO algorithms. Hypotheses have been tested, and the simulation findings have been validated.

The rest of the work is presented as follows. The assumptions and notations are outlined in Section II. A few key concepts for the entire article are presented in Section III. Section IV of the work has explored the development and definition of the System Model. Section V contains the experimental setup and findings. The research's conclusion observations are presented in Section VI and Section VII has discussed Scope for future work.

II. ASSUMPTIONS AND NOTATIONS

The entire paper is based on the following assumptions and notations.

A. ASSUMPTIONS

1. Nodes are arranged inside a square in an even distribution and at random positions.
2. The base station (BS) is the node at which the hierarchical WSN is organized. Between the user and the sensor network, it acts as a communication link. Due to the BS's location outside of the square, it is possible to communicate with nodes subject to multi-path attenuation. Multi-path debilitation has no impact on communication among nodes.
3. Because of their comparable capabilities and initial battery energy levels, the nodes collaborate seamlessly, adapting their functions based on the time of day.
4. The BS and any other node can be reached through communication from any node.
5. Every node remains immobile or stationary.
6. Each node observes its environment and transmits a signal of uniform duration.
7. Parameters in WSNs exhibit uncertainty and are characterized by fuzzy values.

B. NOTATIONS

There are different symbols and notations used in this paper. Table 2 presents a list of symbols used in this paper.

III. PRELIMINARIES

This section offers insights into the utilization of different methods in this research, encompassing the "entropy weighted method", the TOPSIS method involving fuzzy numbers, DBSCAN, and the Silhouette Index.

TABLE 2. List of symbols.

Symbol	Description
dN	Distance between nodes and BS
dN_0	Constant distance measurement to the BS
(CN_x, CN_y)	Co-ordinate of CHs in a Network
(NN_x, NN_y)	Co-ordinate of node in a Network
$EN_{initial}$	Initial resource
EN_{elec}	Electronics energy
EN_{TX}	Energy dissipation in data transmission
εN_{fs}	Energy augmentation to surmount open spaces
εN_{mp}	Energy augmentation for effective multi-path navigation
EN_{RX}	Energy depletion during data reception
N_c	Optimal number of CHs
L	Size of data
N	Nodes in the network overall, in number

A. ENTROPY WEIGHTED METHOD (CF. SEN ET AL. [42])

In our criterion weight calculation, we employed the entropy-weighted technique, which is rooted in the works of Clausius [27] and Shannon [28] for details on entropy. This technique assesses the information-carrying capacity of each evaluation criterion, providing insights into their relative importance. It utilizes entropy values to measure the level of unpredictability in the information. Our initial step in weight computation involves the examination of the decision matrix.

Consider that $Z = (z_{ij})_{m\times n}$ be the decision matrix and $W = (W_1, W_2, \ldots, W_n)$, weights to each criterion, reflecting their relative importance in the decision-making process. These weights should sum to 1 and range in-between [0 1]. Consider a case involving m alternatives $A_i (i = 1, 2, \ldots, m)$ and n criteria $C_j (j = 1, 2, \ldots, n)$. During that period, we can determine the weight W_j, $j = 1, 2, \ldots, n$ by following these steps in the procedure:

Step 1: Calculate $l_{ij} = \frac{z_{ij}}{\sum_{i=1}^{m} z_{ij}}$.

Step 2: Calculate $L_j = -\frac{1}{\log(m)} \sum_{i=1}^{m} l_{ij} \log(l_{ij})$.

It is expected that $l_{ij} \log l_{ij} \to 0$, when $l_{ij} \to 0$.

Step 3: Determine the value of $M_j = 1 - L_j$.

Step 4: Determine the value of $W_j = \frac{M_j}{\sum_{j=1}^{n} M_j} = \frac{1-L_j}{\sum_{j=1}^{n}(1-L_j)}$.

B. TOPSIS METHOD BASED ON TRIANGULAR FUZZY NUMBERS (TFNS) (CF. SEN ET AL. [42])

"TOPSIS (Technique for Order Preference by Similarity to Ideal Solution)" is a "multi-criteria decision-making method". Here, we have incorporated uncertainty in the form of TFNs and we have implemented the TOPSIS method under uncertainty [29], [30], [31] by use of TFNs.

Step 1: Define criteria and alternatives: Suppose a decision-making problem with m criteria $(C_1, C_2, C_3, \ldots, C_m)$ and n alternatives $(A_1, A_2, A_3, \ldots, A_m)$.

Step 2: Triangular Fuzzy Number (TFNs): Each criteria and alternative is represented using TFNs, which are defined by three parameters: a lower bound (a), a modal value (b), and an upper bound (c). A TFN be then $\tilde{z}_{ij} = (a_{ij}, b_{ij}, c_{ij})$

Step 3: Normalization of TFNs: To bring all TFNs to a common scale, normalize each TFN to a crisp value (z_{ij}) using centroid method $z_{ij} = \frac{a_{ij}+2b_{ij}+c_{ij}}{4}$ where:

- (z_{ij}) is the normalized value of the TFN for criterion i, j.
- a_i, b_i and c_i are the lower, modal, and upper values of the TFN for criterion i, respectively.

Step 4: Weighting of Criteria: Assign weights $(W_1, W_2, W_3, \ldots, W_m)$ to each criterion, reflecting their relative importance in the decision-making process. These weights should sum to 1.

Step 5: Normalized Decision Matrix (R): Create a normalized decision matrix R, where each element R_{ij} represents the normalized value of the alternative j for criterion i:

$$R_{ij} = z_{ij} \tag{1}$$

Step 6: Ideal Solution: (A_+): Consider the ideal solution for each criterion, A_+:

$$A_{+i} = 1 \quad i = 1\,(1)\,n$$

where A_{+i} represent ideal solution for criterion i.

Step 7: Non-Ideal Solution: (A_-): Consider the non-ideal solution for each criterion, A_-:

$$A_{-i} = 1 \quad i = 1\,(1)\,n$$

where A_{-i} represent Non-ideal solution for criterion i.

Step 8: Similarity to Ideal Solution (S_+): Calculate the similarity of each alternative to the ideal solution for each criterion.

$$S_{+j} = \sqrt{\sum_{i=1}^{m} (W_i(1 - R_{ij}))^2} \tag{2}$$

Step 9: Similarity to Non-Ideal Solution (S_-): Calculate the similarity of each alternative to the non-ideal solution for each criterion.

$$\mathrm{S_{-j}} = \sqrt{\sum_{\mathrm{i=1}}^{\mathrm{m}} (\mathrm{W_i} \times \mathrm{R_{ij}})^2} \tag{3}$$

Step 10: Relative closeness (C): Calculate the relative closeness for each alternate:

$$C_j = \frac{S_{-j}}{S_{+j} + S_{-j}} \tag{4}$$

Step 11: Rank the alternative based on their relative closeness. The alternative with the highest C_j is considered the most preferred solution.

Equation (4) can determine the most favorable option from a set of predetermined choices and establish a ranking for all available alternatives. Subsequently, these alternatives can be ranked based on their closeness coefficient, with the top-ranked option emerging as the preferred choice.

C. DBSCAN (DENSITY-BASED SPATIAL CLUSTERING)

An established clustering method known as DBSCAN has the ability to detect clusters of varying configurations without the need of number of cluster as input [32]. Epsilon (ε) and minimum points (*Minpts*) are the two required input parameters [33]. *Minpts* is the least number of points within ε radius where ε stands for the neighborhood's radius around any given data point. This algorithm creates a dense cluster of data points and flags the less dense or lone data points as outliers. The following list includes some useful terms associated with DBSCAN:

Eps Neighborhood: Let D is a dataset with any two data points l and m. ε Neighborhood of any point l can be defined as:

$$NEps(l) = \{m \in D \mid d(l, m) \leq \varepsilon\} \tag{5}$$

where $d(l, m)$ denotes distance between data point l and m.

Core point $R_{(cor)}$: When a data point's ε neighborhood exhibits l maximum data points with respect to the *Minpts*, that particular point l is referred to as a core point.

Border point: If the ε vicinity of point l contains a smaller number of data points in contrast to *Minpts*, yet one of the neighboring points is a core point, then point l is designated as a border point.

Noise point: A point l is considered noise or outlier when it can't be categorized by core point or the border point category.

D. SILHOUETTE INDEX (SI)

The Silhouette index (SI) [34] is a well-liked performance statistic used to determine the precision of clustering. It provides the proximity metric for data points contained in the same cluster. Data samples in any clustering pattern exhibit greater similarity to the data in their own cluster and less resemblance to the data in other clusters. The SI value shows high range when similarities between the cluster data are low. SI has a range of $[-1\ 1]$ and larger values are always preferred. In the realm of distributed systems, the SI is expressed in mathematical terms as:

$$SI_c = (\frac{1}{NK}) \sum_{k=1}^{K} \sum_{i=1}^{N} SI(l_i, k) \tag{6}$$

where k epresents sensor node and $k \in [1, K]N$ is total number of data sample at any node. $SI(l_i, k)$ i.e. silhouette index of data point l for $k\ th$ sensor is represented as

$$SI(l, k) = \frac{b(l) - a(l)}{\max\{b(l), a(l)\}} \tag{7}$$

where $a(l)$ indicates the average Euclidean distance of l and other samples in the same class. $b(l)$ s least average Euclidean distance of data sample l to other samples leaving its own cluster.

IV. SYSTEM MODEL DEFINITION AND FORMULATION

The system's infrastructure consists of one base station (BS) and many sensor nodes. All sensor nodes are divided into two groups. The two types of nodes are ordinary nodes and cluster head nodes (CHs). The tasks of ordinary nodes involve overseeing environmental data and relaying sensor data to the CHs. The ordinary nodes are methodically used to select the CH. The CH collects data from regular nodes, which is subsequently transmitted to the BS by the ordinary nodes.

The primary radio model corresponds to the energy paradigm. Here, we solely focus on energy expenditure during the data exchange phase. The complete energy consumption is composed of the energy dissipation arising from data transfer, acquisition, and integration. Within this framework, a data exchange of L-bit occurs between a standard node and a cluster head node, and you can determine energy consumption using the provided equation.

$$EN_{TX}(L, dN) = EN_{elec} * L + \varepsilon N_{amp} * L \tag{8}$$

$$EN_{RX}(L) = EN_{elec} * L \tag{9}$$

$EN_{TX}(L, dN)$ can be defined as the amount of energy used during data dispatch time of data size L-bit and $EN_{RX}(L)$ refers to the energy depletion during the data reception process. You can use Eq. (10) to calculate the energy consumption of the amplifier during the transmission phase, with ε_{amp} denoting the amplifier's energy usage in this stage.

$$\varepsilon N_{amp} = \begin{cases} \varepsilon N_{fs} dN^2 & \text{if } dN \leq dN_0 \\ \varepsilon N_{mp} dN^4 & \text{if } dN > dN_0 \end{cases} \tag{10}$$

If the value of dN greater than dN_0, then the common node will use multipath fading channel which use εN_{fs} and εN_{mp} communication energy parameter or else sensor node will use free-space propagation model. The value of dN_0 can be calculated using Eq.(11).

$$dN_0 = \sqrt{\frac{\varepsilon N_{fs}}{\varepsilon N_{mp}}} \tag{11}$$

V. EXPERIMENTAL SETUP AND COMPUTATIONAL RESULTS

Finding the optimum CH-set that will cover the entire WSNs in the network region is one of the major design issues of a cluster-based routing algorithm [35]. In the context of clustering protocols, it has been observed that during the formation of clusters, each protocol like HEED [36], LEACH-FC and REAC-IN [37] tends to include every nearby node. However, the process of inviting neighboring nodes to join a cluster consumes valuable resources such as propagation time, residual energy and uses of processor etc. This becomes particularly problematic when neighboring nodes are positioned at considerable distances. The actual

Cluster Head can lose its critical resources with each invitation which directly or indirectly effects the overall network lifetime. So there is a need of identification of such kind of noise node. Most of the existing work assumes the presence of CHs in each cluster because having one in a cluster makes maintaining the WSN easier. It is possible to build clusters in a way that should minimize the message overhead that developed during the cluster creation phase. The PSO and BFAO clustering algorithms primarily emphasize optimizing clustering for WSNs. However, the network lifetime is not solely determined by optimized clustering; the selection of appropriate CHs is also a crucial aspect of clustering. In the case of PSO and BFAO, CHs are chosen primarily based on a single parameter, specifically residual energy. Unfortunately, this limited selection criterion leads to frequent cluster formation disruptions, depleting crucial resources. To achieve an optimized cluster configuration, we implemented the DBSCAN based intelligent clustering algorithm. Traditional clustering algorithms like K-Means [38], K-Medoids [39] and hierarchical clustering [40] excel in identifying spherical or convex clusters, making them suitable for compact and well-defined clusters. However, they are sensitive to noise and outliers within the data. The DBSCAN [41] based intelligent algorithm is built on the intuitive concept of distinguishing "clusters" from "noise." In this research, the intelligent clustering algorithm has been applied over 100 nodes, and the optimal number of cluster values has been verified by the elbow method. The DBSCAN-based intelligent clustering has been compared with K-Means-based clustering with different k values. It has been found that the larger value of Silhouette does not always show good clustering. In Table 3, you can find the silhouette scores corresponding to various cluster values.

TABLE 3. Silhouette score for different clusters.

Number of Cluster	2	3	4	5	6	7	8	9
Silhouette Score	0.6905	0.5993	0.6779	0.6120	0.5063	0.4679	0.4240	0.4608

In Table 3, it has been observed that the silhouette score has been varied when the number of clusters has been increased. Cluster value 2, 4, 5,6,7,9 shows a very good silhouette score, but all other clusters leaving cluster 4 have overlapped points depicted by the graph negative values shown in Figure 4-11. The optimized number of cluster values has also been verified by the elbow method, which is 4, shown in Figure 12. The DBSCAN algorithm has been applied to the network, and it has been observed the entire network has been divided into 4 clusters after removing the noise point which is again equal to the elbow score and silhouette score. Figure 13 shows the distribution of nodes after removing noise points using DBSCAN algorithm where $\varepsilon = 0.2$ and $Minpts = 4$.

The selection of accurate CHs is crucial for cluster creation. In this study, CHs are chosen based on various parameters, including (i) Distance from the base station, (ii) Average distance of cluster nodes, (iii) Reliability, and (iv) Residual energy. However, the influence of each parameter on the CHs selection is not uniform. For this reason entropy weighted technique has been employed to determine the weight of each parameter in CHs selection, which provides a measure of the information content. If we consider Distance from the base station as (α), Average distance of cluster nodes

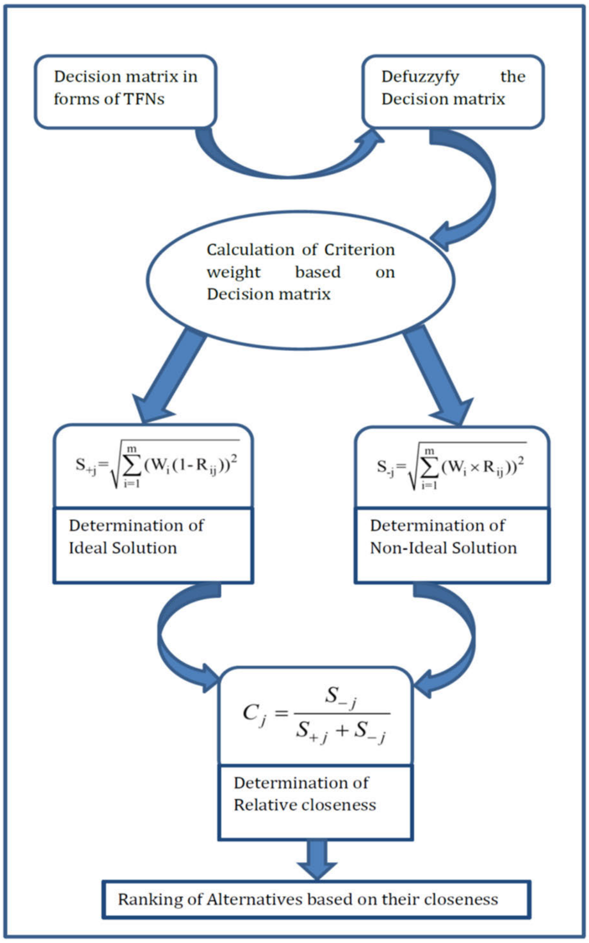

FIGURE 3. Figure shows the steps to finds the rank of alternative solution.

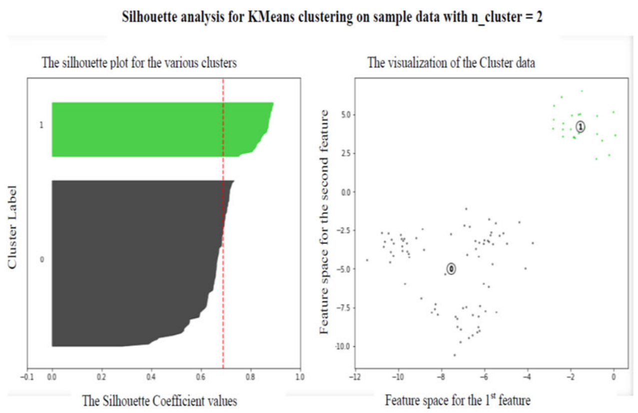

FIGURE 4. The distribution of data with 2 clusters along with Silhouette analysis.

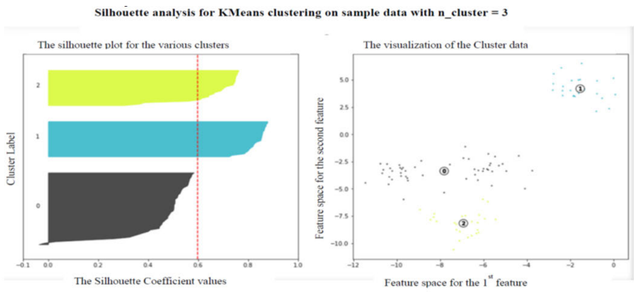

FIGURE 5. The distribution of data with 3 clusters along with silhouette analysis.

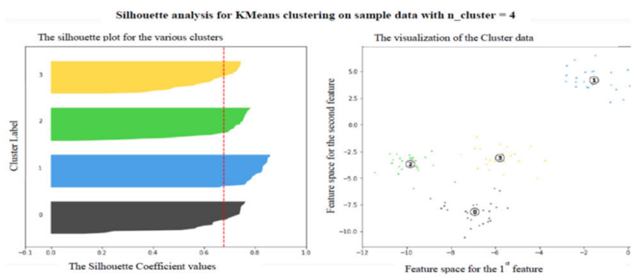

FIGURE 6. Distribution of data with 4 clusters along with silhouette analysis.

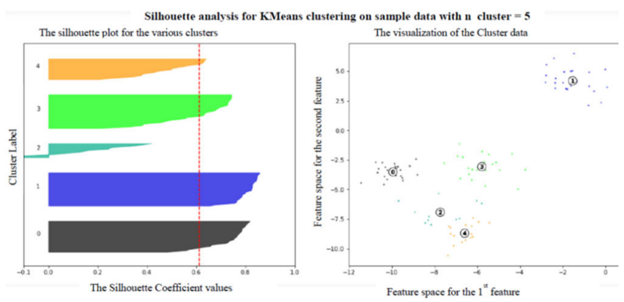

FIGURE 7. The distribution of data with 5 clusters along with silhouette analysis.

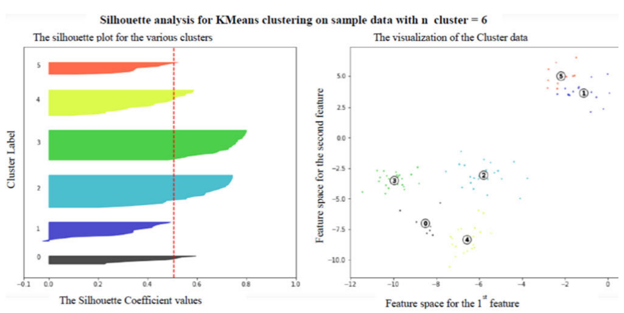

FIGURE 8. The distribution of data with 6 clusters along with silhouette analysis.

(β), Reliability (γ) and Residual energy (λ) then weight of *Node i* can be calculated as:

$$Node\,(i) = (0.1060/\alpha) + (0.1215/\beta) + (0.3645^{*}\gamma) + (0.4078^{*}\lambda) \quad (12)$$

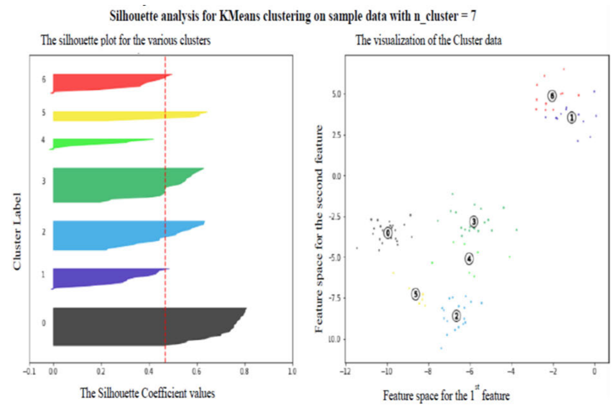

FIGURE 9. The distribution of data with 7 clusters along with silhouette analysis.

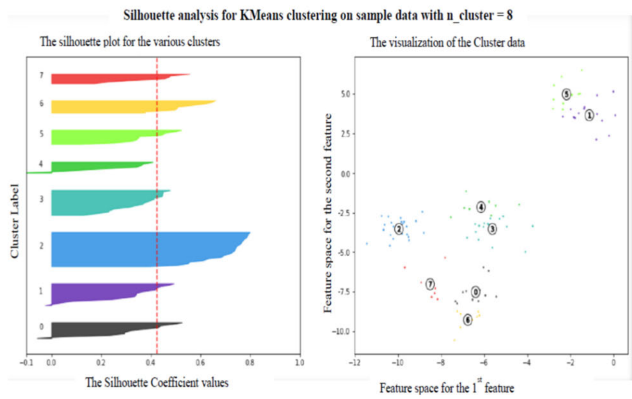

FIGURE 10. The distribution of data with 8 clusters along with silhouette analysis.

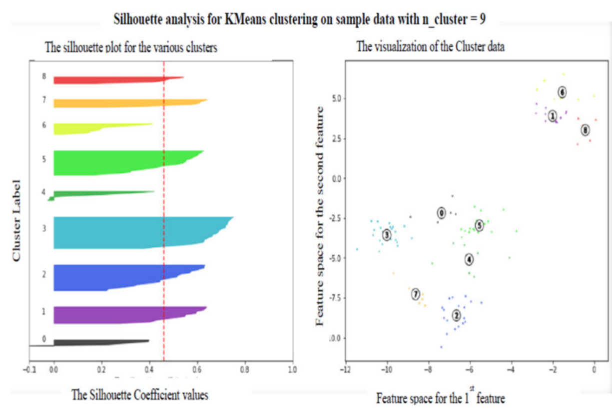

FIGURE 11. The distribution of data with 9 clusters along with silhouette analysis.

We utilized the entropy method to determine the weight assigned to each feature or characteristic, as illustrated in Table 5. The TOPSIS are mainly used for comprehensive decision-making. The advantages of TOPSIS for the selection of cluster heads in WSNs lie in its ability to handle multi-criteria decision problems, balance conflicting objectives, offer transparency in decision-making and adapt to

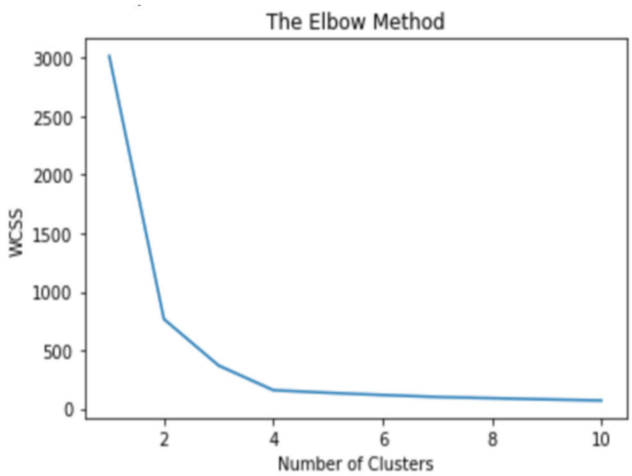

FIGURE 12. The optimal number of clusters using elbow method.

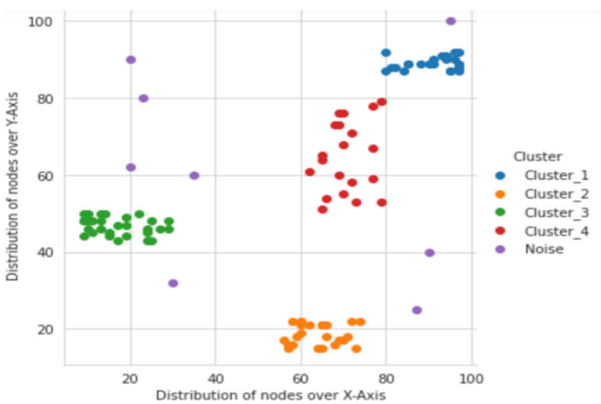

FIGURE 13. Distribution of 100 nodes using DBSCAN based intelligent clustering algorithm.

real-world uncertainties for improved network performance. In Cluster 3, it has been observed that two nodes had identical weights calculated using the entropy-weighted technique. Nodes with identical weights cannot be selected as CHs. To address this issue, the TOPSIS technique is employed, assigning a rank to each node within the cluster. Table 4 displays the rank of the top 10 nodes in Cluster 3. Despite having the same entropy-weighted weight, Node 5 and Node 2 obtained different ranks through the TOPSIS method. Consequently, based on the rank, Node 5 is selected as the CH.

TABLE 4. Rank of top 10 nodes with entropy weighted values.

Rank(i)	1	2	3	4	5	6	7	8	9	10
Node	Node5	Node2	Node3	Node9	Node10	Node8	Node7	Node 14	Node4	Node 15
Weight	0.506	0.506	0.501	0.493	0.482	0.476	0.462	0.451	0.449	0.436

MCDM technique has been applied to each cluster to select CH from each cluster. Table 6 shows the CHs of each cluster along with its parameter values.

TABLE 5. Entropy weighted value of each feature/characteristic.

Properties	Weight
Distance from base station	0.1060
Average distance of cluster nodes	0.1215
Reliability	0.3645
Residual energy	0.4078

TABLE 6. Four cluster head along with their all-parametric values.

Cluster Head	Distance from Sink	Average Distance of cluster nodes	Reliability of Cluster	Residual Energy
CH 0	75.480	7.364	0.310	0.9698
CH 1	86.281	6.987	0.215	0.9657
CH 2	67.242	9.239	0.472	0.9798
CH 3	65.208	8.891	0.268	0.9648

In our investigation, we created a network of 100 nodes, with the BS and nodes spread randomly around the area. Each data message consisted of 4000 bits, while the packet header for each packet type had a size of 25 bytes. The channel's data transmission rate was configured to be 1 megabyte per second. We utilized the optimal cluster count to ascertain the initial number of clusters. As inactive nodes were detected, the cluster count was dynamically changed according to the density of active nodes, resulting in the consolidation of smaller clusters with larger ones. For this study, we assumed that BS nodes had advanced processing capabilities and had access to a plentiful energy source.

For the interest of experiments, we have determined the optimum cluster number which is 4. Here M $=$ 100 m, N $=$ 100 nodes, $\varepsilon_{fs} = 10$ pJ and ε_{mp}=0.001275pJ.

Furthermore, we have taken into account the unforeseen initial energy levels of each node and the unpredictable data packet sizes, which are represented as (0.7, 1, 1.2) and (495, 500, 510) in a TFN. Table 7 shows all the parameter along with its values which are used for experiments.

TABLE 7. Defuzzified values of all parameters.

Parameters	Parametric value as per assumptions	Defuzzified value
N	100	
$EN_{initial}$	(0.7,1,1.2) J/bit/m^2	0.975 J/bit/m^2
Coordinate of BS	(50,100)	
Size of the data packet	(495,500,510) byte	501.25 byte
Hello/broadcast/CH join message	(22,25,28) byte	25 byte
εN_{fs}	(8,10,12) pJ/bit/m^2	10 pJ/bit/m^2
εN_{mp}	(0.001,0.0013,0.0015) pJ/bit/m^2	0.001275 pJ/bit/m^2
EN_{elec}	(47,50,52) pJ/bit/m^2	49.75 pJ/bit/m^2
Optimal Cluster	4	

A. NODE SELECTION CRITERIA

Initially, the DBSCAN algorithm was used to partition the entire network into clusters. The elbow technique and the Silhouette index have been used to determine the optimal number of clusters. With these two methods, the cluster value of the DBSCAN algorithm has been verified. By applying the entropy technique, various weights are provided for each parameter. To choose the best CHs for simulation in the first round, the TOPSIS approach has been used. To identify CHs for subsequent rounds, we have proposed an algorithm. Over 100 nodes, at first, we have chosen the best CHs using the DBSCAN based intelligent clustering. Figure 14 shows the steps for selection of CHs through intelligent clustering using entropy weighted MCDM technique.

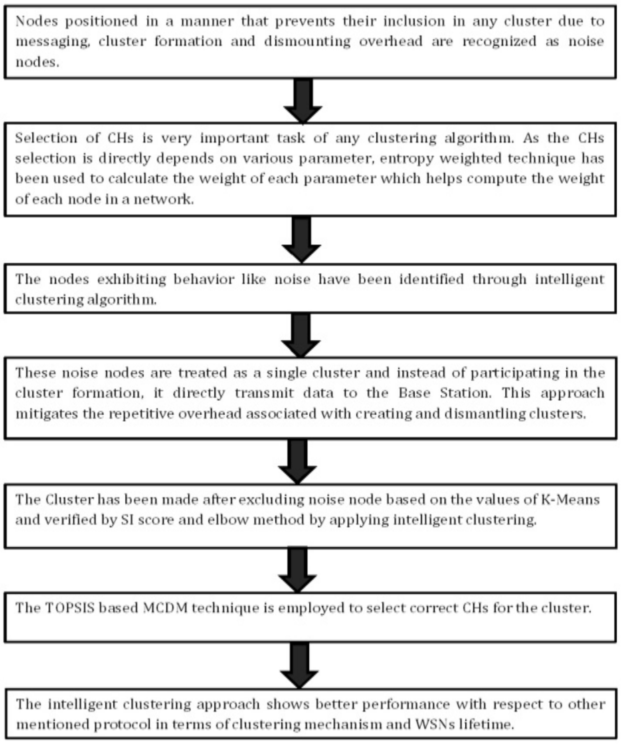

FIGURE 14. **Steps for creation of cluster and CHs via intelligent clustering.**

The CH selection algorithm has been implemented in Python and executed in a Python Jupyter Notebook (Version 3) on a Linux platform. To determine the CHs for the upcoming rounds, we have applied our proposed Algorithm 1, which considers multiple criteria such as the average distance among cluster nodes, residual energy, reliability, and distance from the sink. Algorithm 1 is as follows:

Algorithm 1:

Step 1: The initial optimal number of clusters has been obtained by SI score, verified by the elbow method.

Step 2: Clustering has been done using the DBSACN intelligent clustering algorithm, verifying the optimal clustering value calculated by the SI score.

Step 3: Each node has sent its data to the BS for the first simulation round.

Step 4: Calculate the values of average Distance of cluster nodes, reliability, residual energy, and distance from the sink of each node. The term "average Distance of cluster nodes" means average distance between the node and its neighbor nodes with in a cluster. If dN be the distance between CH and a designated node then dN can be measured using the formula

$$dN = \sqrt{(CN_x - NN_x)^2 + (CN_y - NN_y)^2}.$$

Reliability is calculated by

$$1 - \frac{\text{No. of Neighbors in a cluster}}{\text{Maximum number of Nodes in a cluster} + 1}.$$

Step 5: TOPSIS and the Entropy technique has been used to find the CHs for respective clusters.

Step 6: Stop.

B. SIMULATION RESULTS AND DISCUSSIONS

We have proposed and developed another algorithm for simulation purposes. We refer to it as Algorithm 2 here. Algorithm 2 is as follows:

Algorithm 2:

Step 1: 100 nodes have been deployed randomly over (100, 100)m^2 area with BS (50, 100) oordinates.

Step 2: For the Second and subsequent rounds, the selected CHs will send the data which has been selected by applying Algorithm 1.

Step 3: Repeat Steps 4 to 9 until all nodes' residual energy is not diminished.

Step 4: Increase a counter whenever the residual energy of a node surpasses that of every other node in its respective cluster.

Step 5: Increment a counter when a node's distance from the sink surpasses that of all other nodes in the cluster.

Step 6: Increment a counter when a node's average distance from other nodes in the cluster is greater than that of all other nodes.

Step 7: The node with the highest counter value has been considered CHs for the next round.

Step 8: If a cluster contains less than three nodes, assign additional nodes to the nearest cluster based on the reliability of each cluster.

Step 9: Move to the next round.

Step 10: End.

Algorithm 2 has been implemented using C++ programming, and the graphical presentation has been done using Python programming based on the numerical results produced by Algorithm 2. After 3800 rounds, the residual energy of each node has diminished (gone down), shown clearly in Figure 15. The comparable situation where the residual energy needs to be reduced to use the clustering method is shown by the dotted line in Figure 15. Figure 15 shows that the residual energy fully diminished after 3800cycles.

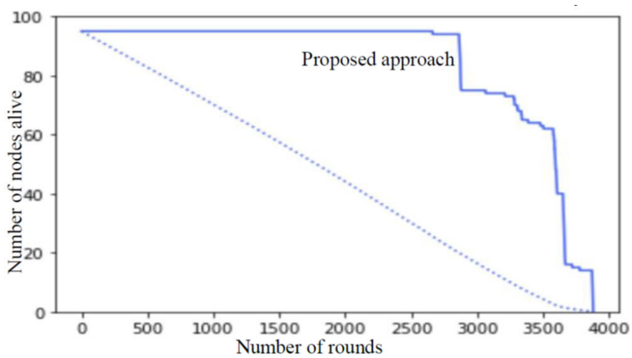

FIGURE 15. **Number of nodes alive vs. number of rounds using proposed approach.**

The simulation was executed 100 times, The expected lifetime of each cluster is displayed as a line graph in Figure 16.

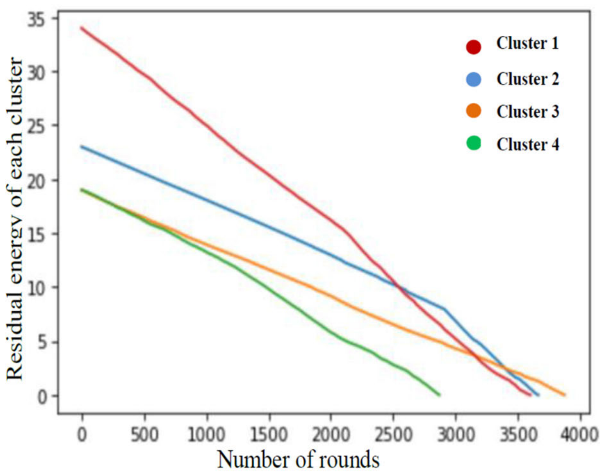

FIGURE 16. **Lifetime of each cluster through line graph.**

Energy usage is one of the major challenges when WSN design is a concern. WSNs are constructed with limited energy. There are several algorithms, like: HEED which a hybrid approach that includes both centralized and distributed mechanism for cluster formation. Mathematically the probability P_i of node i becoming a cluster head can be expressed as $P_i = \frac{G_i}{\sum_{j=1}^{n} G_j}$ where G_i represents the energy of node i, and N is the total number of nodes in the network. It has been observed that in HEED the selection of CHs totally based on probability and any weakest node also can be selected as CHs which may lead a problem for cluster creation. REAC-IN clustering refers to clustering algorithm in WSNs that take into consideration the remaining energy levels of sensor nodes when forming andmaintaining cluster. The energy balance metric of energy balance clustering can be formulated as

$$EnergyBalanceMetric = \frac{S\tan dred\ Deviation\ of\ Energy\ Levels}{Average\ Energy\ Levels}.$$

In REAC-IN formation of cluster and selection of CHs are based on single parameter which is residual energy. Several studies [43], [44], [45] have highlighted the inadequacy of relying on a single parameter for CH selection, leading to repeated disruptions in cluster formation. In this research, instead of choosing probabilistic methods or single-parameter selection for CHs, CHs were chosen based on multiple parameters resulting in a more precise selection during the cluster formation process.

The LEACH Fuzzy clustering protocol, adopting a centralized approach, selects CHs based on (i) Node energy, (ii) Node concentration, and (iii) Node centrality. In contrast, optimizing algorithms such as BFAO and PSO primarily focus on optimize clustering without addressing the selection of appropriate CHs. Both approaches have limitations concerning the extension of WSNs lifetime.

In our proposed approach, CHs are chosen through an entropy-weighted criteria and MCDM technique. Optimal clustering is achieved using the K-Mean algorithm, validated through the elbow method and SI index. Intelligent clustering has been done using DBSCAN algorithm. The results demonstrate an improvement in WSNs lifetime compared to LEACH-FC, PSO, and BFAO-based approaches. A comparison has been made between REAC-IN, LEACH-FC and HEED with the proposed approach (see Fig. 17). The suggested approach outperforms REAC-IN, LEACH-FC and HEED by showing 38%, 15% and 43% more network lifetime.

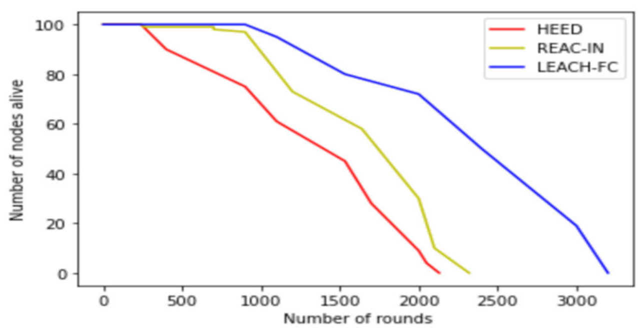

FIGURE 17. **Number of rounds vs. number of nodes alive using the HEED, REAC-IN and LEACH-FC algorithms.**

The clustering strategy mainly focuses on ensuring a high lifetime. There are several possible ways to create effective clustering [46], [47]. The optimization method is one of them, which ensures that the number of clusters should be optimized [48], [49]. The optimized number of clusters is based on inter and intra-cluster distance values. Further studies have been done to find an optimal number of clusters, and it has been found that several optimization techniques like OCABC, PSO and BFAO have been implemented to find an optimal number of clusters. This study's experimental result has also been compared with PSO and BFAO-applied clustering problems in WSN. It has been found that the proposed approach shows 35% and 22% more network lifetime as compared with PSO and BFAO optimization techniques (see Figure 18). Once more, the clustering strategy for WSNs enhances the network's longevity and energy efficiency. Unfortunately, the network may not last as long if a cluster only has one node or if the clustering algorithm chooses nodes close together as CHs for different clusters. Consequently, the choice of CHs holds the utmost significance. Several existing approaches completely disregard factors such as average distance and residual energy in their CH selection process. In our proposed algorithm, we primarily consider the following factors for CH selection: distance from the base station, average distance between cluster nodes, and residual energy. Additionally, we address the issue of clusters with a single node by incorporating a reliability parameter into our approach.

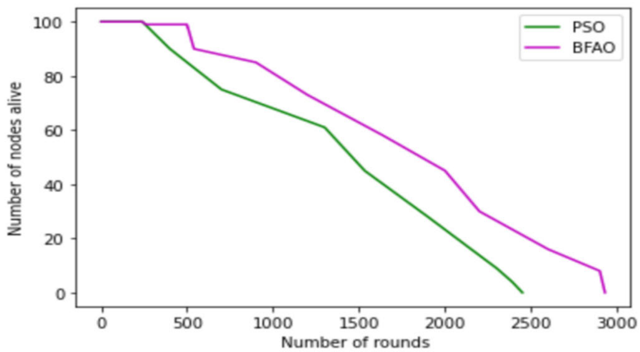

FIGURE 18. Number of rounds vs. number of nodes alive using PSO and BFAO algorithms.

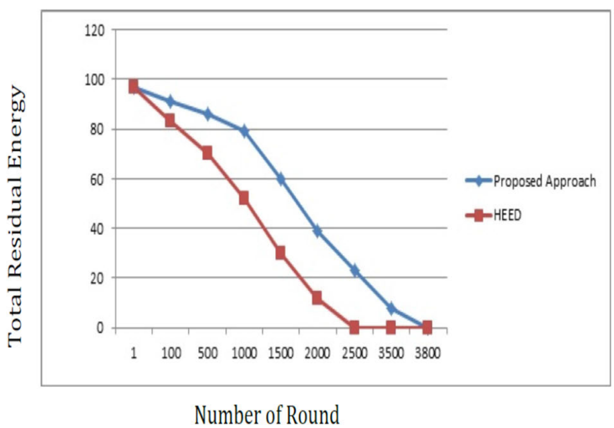

FIGURE 19. Total residual energy of nodes vs. number of round of proposed approach and HEED.

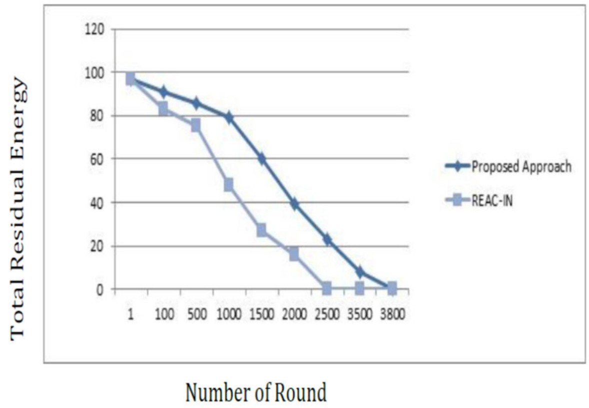

FIGURE 20. Total residual energy of nodes vs. number of round of proposed approach and REAC-IN.

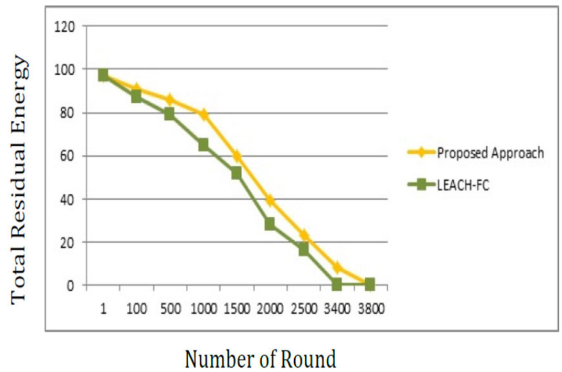

FIGURE 21. Total residual energy of nodes vs. number of round of Proposed approach and LEACH-FC.

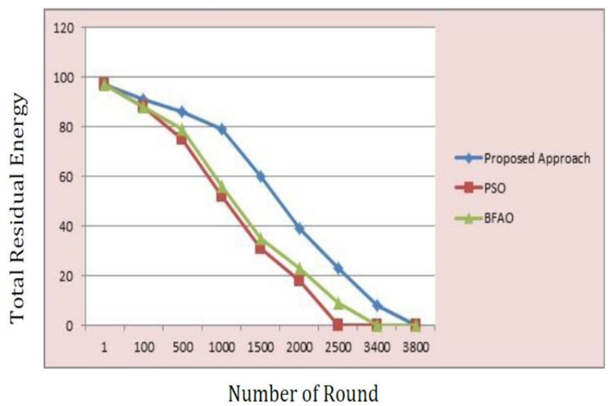

FIGURE 22. Total residual energy of nodes vs. number of round of proposed approach, PSO and BFAO.

In this study Figure 19- Figure 22 shows the comparison of total residual energy and number of rounds of proposed Approach against HEED, REAC-IN, LEACH-FC, PSO and BFAO clustering algorithm. The study has found that the proposed approached outperform the algorithm by 43% (HEED), 38% (REAC-IN), 15% (LEACH-FC), 35 %(PSO) and 22 %(BFAO).

Proper clustering and selection of right CHs can avoids repeatedly creation and breaking of clustering. Figure 23 shows that in proposed approach first re-clustering have been done after 2800 number of round where as in HEED, REAC-IN, LEACH-FC, PSO and BFAO has perform re-clustering after 200, 800, 1300, 400, 450 number of rounds.

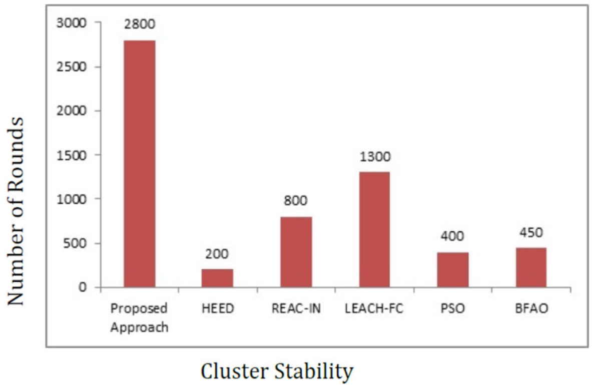

FIGURE 23. Stability of network.

Computational Complexity: Take N_c to represent the optimal cluster count and N as the total node count. Evaluating the appropriate cluster count N_c requires $\Theta(N_c)$ iterations. Computing the average distance between each pair of nodes within a cluster involves $\Theta(N^2)$ calculations. The computation of residual energy and distance from the sink can be completed in constant time, referred to as ψ_1.

Hence, the time complexity of the algorithm can be represented as $\Theta(N_cN^2 + \psi_1) \approx \Theta(N_cN^2)$.

Nevertheless, in situations where the number of nodes and clusters is equal, the worst-case time complexity increases to $\Theta(N^3)$.

Validity: The t-test is a statistical hypothesis test used to determine if there is a significant difference between the means of two groups. It is commonly used when dealing with small sample sizes. The t-test can be one-sample, independent samples, or paired samples, depending on the experimental design. A mathematical explanation of t-test with given significance level given below

T-Test:

Null Hypothesis (H_0): There is no significant difference between the sample mean ($\widetilde{X}$) and a known or hypothesized population mean (μ).

Alternative Hypothesis (H_1)**:** There is a significant difference between the sample mean ($\widetilde{X}$) and a known or hypothesized population mean (μ).

Test Statistic:

$$t = \frac{\widetilde{X} - \mu}{\left(\frac{S}{\sqrt{n}}\right)}$$

where:

- $\widetilde{X}$ is the sample mean.
- μ is the population mean.
- S is the sample standard deviation.
- n is the sample size.

Degrees of Freedom (df)

$$df = n - 1$$

Decision Rule:

If the absolute value of t is greater than the critical t-value for the chosen significance level, then the Null hypothesis is rejected.

Confidence Interval (CI) for the Mean:

$$CI = \widetilde{x} \pm t_{\frac{\alpha}{2}}\left(\frac{S}{\sqrt{n}}\right)$$

where:

$t_{\alpha/2}$ Is the critical t-value for the chosen significance level $\left({}^{\alpha}/_{2}\right)$ and degree of freedom (df)**.**

In this study the result has been verified with t test statistic.

Null Hypothesis (H_0): There is no significant difference between the sample average number of rounds ($\widetilde{X}$) and actual average number of rounds (μ).

Alternative Hypothesis (H_1)**:** There is a significant difference between the sample average number of rounds ($\widetilde{X}$) and actual average number of rounds (μ).

Test Statistic:

Now, $\widetilde{X} = 3805, \mu = 3793, S = 19.5, n = 10,$
$t = 1.92$ and $CI = [3793, 3817]$.

For a two-tailed test at a 95% confidence interval, the t-value is approximately 2.093.

Since$|t| = 1.92 < 2.093$, there is no evidence to reject null hypothesis.so, there is no significant difference between the sample average number of rounds ($\widetilde{X}$) and actual average number of rounds (μ).

VI. CONCLUDING REMARKS

We used multi-criteria decision-making and DBSCAN-based intelligent clustering approaches in this study to choose the best CHs that would maximize network lifetime when there was uncertainty. Designing WSNs which is adaptable to diverse environments requires robust protocols and algorithms that can dynamically adjust to changing conditions. Uncertainties in environmental factors can affect the accuracy of data and reliability of cluster. In this study the uncertainty of data is handled by TFNs along with the reliability of cluster. We have introduced a novel simulation algorithm aimed at improving network longevity in the environment of uncertainty. As digital content experiences exponential growth and relies on network channels for transmission, WSNs play a crucial role in communication. But making networks last longer is hard, and we need effective routing algorithms and data compression methods that use little energy, like image and video compression. In order to tackle these challenges, the papers introduce a clustering algorithm that relies on DBSCAN and employs the K-Means and elbow methods to determine the ideal cluster count. These strategies strive to minimize energy usage by optimizing the utilization of remaining energy and improving the overall performance of the network. Finding the ideal number of clusters is a challenging endeavor that involves optimizing the distances between nodes both within and between clusters. This optimization challenge can be addressed as either a single-objective or a multi-objective problem. Popular optimization models for multiple-objective problems include PSO and BFAO. While these models have been extensively studied and proven to improve network lifetime dramatically, they need to improve their execution times and algorithmic complexity compared to the proposed approach.

Deploying WSNs in real-world scenarios involves strategically placing nodes, a task influenced by various environmental constraints. Intelligent clustering identifies nodes positioned in locations incompatible with cluster formation. This process reduces clustering overhead, contributing to an extended WSNs lifetime.

As WSNs often require scaling to accommodate increased nodes or expanded geographical coverage, intelligent clustering becomes crucial. In scenarios involving the adoption of new nodes the intelligent clustering automatically integrates nodes into existing clusters or identifies them as noise nodes based on their positions. This adaptive approach enhances network lifetime, demonstrating the system's resilience and scalability. The research conducted in this paper includes a comparative analysis between the proposed approach, PSO, and BFAO in terms of network lifetime. The results indicate that the proposed approach outperforms BFAO and PSO by achieving a 22% and 35% increase in network lifetime, respectively. Current research into extending network lifetimes predominantly centers on optimal clustering and CH selection. Conversely, the proposed method integrates optimal clustering through the utilization of K-Means and the elbow method alongside CH selection. In the past, studies that used the PSO and BFAO algorithms only looked at optimal clustering and did not look at how to pick the best CHs from each cluster.

Consequently, the proposed approach demonstrates improved cluster stability over multiple simulations compared to the alternatives. Diagrams depicting this behavior are presented in Fig. 15 and Fig. 16. Experimental findings and analysis of the suggested scheme illustrate that the proposed approach significantly reduces energy consumption and extends network lifetime compared to well-known algorithms such as LEACHFC, REAC-IN and HEED. It is worth noting that the research assumes a two-dimensional node position, which does not reflect the reality of multi-dimensional node positioning. This oversight may impact the overall network lifetime.

This study incorporates intelligent clustering to identify noise nodes, which create challenges during cluster creation. If such nodes are absent, intelligent clustering becomes unnecessary, which will eliminate additional overhead for cluster creation. LEACH-FC excels in selecting appropriate CHs based on various parameters and incorporates fuzzy clustering, potentially outperforming the proposed approach in the absence of noise nodes. On the other hand, BFAO and PSO-based clustering primarily concentrate on optimizing clustering without addressing CH selection. In scenarios without noise nodes, if PSO and BFAO consider additional parameters such as distance from the base station and average distance of non-cluster nodes during CH selection, these clustering algorithms may surpass the proposed approach in the context of network lifetime.

VII. FUTURE SCOPE FOR RESEARCH

Selection of right CHs is depends on various properties. In this research CHs has been selected based on the criteria's like (i) Distance from base station (ii) Average distance of cluster nodes (iii) Reliability and (iv) Residual energy. There are other properties like signal to noise ratio (SNR), data accuracy and data security etc. can also be suggested as section criteria of CHs. The simulation has been done based on 2D axis, but in real world scenario this result might not be effective. If one can perform the simulation in multidimensional environment then the result will be more acceptable. There are several other ways to represent fuzzy like (i) Interval environment and (ii) Stochastic environment. Future studies can be made on representing uncertainty through Interval environment and stochastic environment. Our proposed algorithm can take $\Theta(N^3)$ time in worst scenario. Further research can be made to improve the execution time of the proposed algorithm. Additionally, the distance calculation between nodes does not consider obstacles, whereas real-life scenarios involve numerous obstacles that can diminish the algorithm's effectiveness. Further studies should account for obstacles when calculating node distances and consider the energy consumed during data transmission across the network. The suggested approach shows potential for future studies in the identification of CHs through multi-criteria decision-making, thereby potentially extending the lifetime of networks. Moreover, this approach can address various WSN problems by incorporating different parameters and models, offering opportunities for further research and exploration.

AUTHORSHIP CONTRIBUTION STATEMENT

"All authors listed have significantly contributed to the development and the writing of the article."

DECLARATION OF COMPETING INTEREST

"The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper."

DATA AVAILABILITY

Not applicable.

ACKNOWLEDGMENT

The authors would like to thank the contributions of the different team members of the VOTE-TRA project (https://www.sfi.ie/challenges/digital-for-resilience/VOTE_TRA/.

REFERENCES

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: A survey," *Comput. Netw.*, vol. 38, no. 4, pp. 393–422, 2002.

[2] R. Wazirali, R. Ahmad, A. Al-Amayreh, M. Al-Madi, and A. Khalifeh, "Secure watermarking schemes and their approaches in the IoT technology: An overview," *IEEE Wireless Commun.*, vol. 15, pp. 60–66, 2021.

[3] H. Khattak, Z. Ameer, U. Din, and M. Khan, "Cross-layer design and optimization techniques in wireless multimedia sensor networks for smart cities," *Comput. Sci. Inf. Syst.*, vol. 16, no. 1, pp. 1–17, 2019.

[4] K. Haseeb, I. Ud Din, A. Almogren, N. Islam, and A. Altameem, "RTS: A robust and trusted scheme for IoT-based mobile wireless mesh networks," *IEEE Access*, vol. 8, pp. 68379–68390, 2020.

[5] A. Banerjee, D. Garg, V. Das, L. Sahoo, I. Nath, V. Varadarajan, and K. Kotecha, "Design of energy efficient WSN using a noble SMOWA algorithm," *Comput., Mater. Continua*, vol. 72, no. 2, pp. 3585–3600, 2022.

[6] H. Kim and S.-W. Han, "An efficient sensor deployment scheme for large-scale wireless sensor networks," *IEEE Commun. Lett.*, vol. 19, no. 1, pp. 98–101, Jan. 2015.

[7] A. Ali, Y. Ming, S. Chakraborty, and S. Iram, "A comprehensive survey on real-time applications of WSN," *Future Internet*, vol. 9, no. 4, p. 77, Nov. 2017.

[8] Z. Ahmad, A. S. Khan, K. Nisar, I. Haider, R. Hassan, M. R. Haque, and J. J. Rodrigues, "Anomaly detection using deep neural network for IoT architecture," *Appl. Sci.*, vol. 11, no. 15, p. 7050, 2021.

[9] L. Sahoo, S. Sen, K. Tiwary, S. Samanta, and T. Senapati, "Modified Floyd–Warshall's algorithm for maximum connectivity in wireless sensor networks under uncertainty," *Discrete Dyn. Nature Soc.*, vol. 2022, pp. 1–15, 2022.

[10] L. Sahoo, S. Sen, K. Tiwary, S. Samanta, and T. Senapati, "Optimization of data distributed network system under uncertainty," *Discrete Dyn. Nature Soc.*, vol. 2022, pp. 1–12, Apr. 2022.

[11] C. Gherbi, Z. Aliouat, and M. Benmohammed, "An adaptive clustering approach to dynamic load balancing and energy efficiency in wireless sensor networks," *Energy*, vol. 114, pp. 647–662, Nov. 2016.

[12] W. Zou, Y. Zhu, H. Chen, and X. Sui, "A clustering approach using cooperative artificial bee colony algorithm," *Discrete Dyn. Nature Soc.*, vol. 2010, Nov. 2010, Art. no. 459796.

[13] P. Vora and B. Oza, "A survey on K-mean clustering and particle swarm optimization," *Int. J. Sci. Modern Eng.*, vol. 1, no. 3, pp. 24–26, 2013.

[14] S. R. Deepa and D. Rekha, "Bacterial foraging optimization-based clustering in wireless sensor network by preventing left-out nodes," in *Intelligent Computing Paradigm: Recent Trends*, 2020, pp. 43–58.

[15] P. Dutta, "Multi-criteria decision making under uncertainty via the operations of generalized intuitionistic fuzzy numbers," *Granular Comput.*, vol. 6, no. 2, pp. 321–337, Apr. 2021.

[16] A. C. Tolga, I. B. Parlak, and O. Castillo, "Finite-interval-valued type-2 Gaussian fuzzy numbers applied to fuzzy TODIM in a healthcare problem," *Eng. Appl. Artif. Intell.*, vol. 87, Jan. 2020, Art. no. 103352.

[17] J. R. Castro, M. A. Sanchez, C. I. Gonzalez, P. Melin, and O. Castillo, "A new method for parameterization of general type-2 fuzzy sets," *Fuzzy Inf. Eng.*, vol. 10, no. 1, pp. 31–57, Jan. 2018.

[18] J.-S. Leu, T.-H. Chiang, M.-C. Yu, and K.-W. Su, "Energy efficient clustering scheme for prolonging the lifetime of wireless sensor network with isolated nodes," *IEEE Commun. Lett.*, vol. 19, no. 2, pp. 259–262, Feb. 2015.

[19] B. Solaiman and A. Sheta, "Computational intelligence for wireless sensor networks: Applications and clustering algorithms," *Int. J. Comput. Appl.*, vol. 73, no. 15, pp. 1–8, Jul. 2013.

[20] J. RejinaParvin and C. Vasanthanayaki, "Particle swarm optimization-based clustering by preventing residual nodes in wireless sensor networks," *IEEE Sensors J.*, vol. 15, no. 8, pp. 4264–4274, Aug. 2015.

[21] E. Moharamkhani, B. Zadmehr, S. Memarian, M. J. Saber, and M. Shokouhifar, "Multiobjective fuzzy knowledge-based bacterial foraging optimization for congestion control in clustered wireless sensor networks," *Int. J. Commun. Syst.*, vol. 34, no. 16, p. 4949, Nov. 2021.

[22] T. Senthil Murugan and A. Sarkar, "Optimal cluster head selection by hybridisation of firefly and grey wolf optimisation," *Int. J. Wireless Mobile Comput.*, vol. 14, no. 3, p. 296, 2018.

[23] D. Wohwe Sambo, B. Yenke, A. Förster, and P. Dayang, "Optimized clustering algorithms for large wireless sensor networks: A review," *Sensors*, vol. 19, no. 2, p. 322, Jan. 2019.

[24] S. Lata, S. Mehfuz, S. Urooj, and F. Alrowais, "Fuzzy clustering algorithm for enhancing reliability and network lifetime of wireless sensor networks," *IEEE Access*, vol. 8, pp. 66013–66024, 2020.

[25] Y. U. Xiu-wu, Y. U. Hao, L. Yong, and X. Ren-rong, "A clustering routing algorithm based on wolf pack algorithm for heterogeneous wireless sensor networks," *Comput. Netw.*, vol. 167, Feb. 2020, Art. no. 106994.

[26] S. R. Deepa and D. Rekha, "Cluster optimization in wireless sensor networks using particle swarm optimization," in *Proc. 52nd Annu. Conv. Comput. Society India*, Kolkata, India, Singapore: Springer, 2017, pp. 240–253.

[27] R. Clausius, "I. On the moving force of heat, and the laws regarding the nature of heat itself which are deducible therefrom," *London, Edinburgh, Dublin Phil. Mag. J. Sci.*, vol. 2, no. 8, pp. 1–21, Jul. 1851.

[28] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, Oct. 1948.

[29] C. H. Chen, "A novel multi-criteria decision-making model for building material supplier selection based on entropy-AHP weighted TOPSIS," *Entropy*, vol. 22, no. 2, p. 259, 2020.

[30] L. Sahoo, "Some score functions on fermatean fuzzy sets and its application to bride selection based on TOPSIS method," *Int. J. Fuzzy Syst. Appl.*, vol. 10, no. 3, pp. 18–29, Jul. 2021.

[31] J.-S. Yao and K. Wu, "Ranking fuzzy numbers based on decomposition principle and signed distance," *Fuzzy Sets Syst.*, vol. 116, no. 2, pp. 275–288, Dec. 2000.

[32] S. J. Nanda and G. Panda, "Design of computationally efficient density-based clustering algorithms," *Data Knowl. Eng.*, vol. 95, pp. 23–38, Jan. 2015.

[33] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proc. KDD*, 1996, vol. 96, no. 34, pp. 226–231.

[34] D. K. Kotary and S. J. Nanda, "Distributed robust data clustering in wireless sensor networks using diffusion moth flame optimization," *Eng. Appl. Artif. Intell.*, vol. 87, Jan. 2020, Art. no. 103342.

[35] S. F. Hussain and S. Iqbal, "CCGA: Co-similarity based co-clustering using genetic algorithm," *Appl. Soft Comput.*, vol. 72, pp. 30–42, Nov. 2018.

[36] P. Gupta and A. K. Sharma, "Clustering-based heterogeneous optimized-HEED protocols for WSNs," *Soft Comput.*, vol. 24, no. 3, pp. 1737–1761, Feb. 2020.

[37] K. B. Gurumoorthy, S. A. Devaraj, S. Gopinath, and T. Ali, "A novel clustering method for fault recovery and routing in mobile ad-hoc networks," *Int. J. Commun. Syst.*, vol. 34, no. 15, p. 4937, 2021.

[38] M. Sathyamoorthy, S. Kuppusamy, R. K. Dhanaraj, and V. Ravi, "Improved K-means based Q learning algorithm for optimal clustering and node balancing in WSN," *Wireless Pers. Commun.*, vol. 122, no. 3, pp. 2745–2766, Feb. 2022.

[39] S. Zhang, Y. Wang, Y. Zhang, P. Wan, and J. Zhuang, "Riemannian distance-based fast K-medoids clustering algorithm for cooperative spectrum sensing," *IEEE Syst. J.*, vol. 16, no. 1, pp. 880–890, Mar. 2022.

[40] S. Zafar, A. Bashir, and S. A. Chaudhry, "Mobility-aware hierarchical clustering in mobile wireless sensor networks," *IEEE Access*, vol. 7, pp. 20394–20403, 2019.

[41] M. K. Hameed and A. K. Idrees, "Distributed DBSCAN protocol for energy saving in IoT networks," in *Proc. Int. Conf. Commun., Comput. Electron. Syst. (ICCCES)*. Singapore: Springer, 2020, pp. 11–24.

[42] S. Sen, L. Sahoo, K. Tiwary, V. Simic, and T. Senapati, "Wireless sensor network lifetime extension via K-medoids and MCDM techniques in uncertain environment," *Appl. Sci.*, vol. 13, no. 5, p. 3196, Mar. 2023.

[43] P. Mukherjee, P. K. Pattnaik, A. A. Al-Absi, and D.-K. Kang, "Recommended system for cluster head selection in a remote sensor cloud environment using the fuzzy-based multi-criteria decision-making technique," *Sustainability*, vol. 13, no. 19, p. 10579, Sep. 2021.

[44] B. Raj, I. Ahmedy, M. Y. I. Idris, and R. Md. Noor, "A survey on cluster head selection and cluster formation methods in wireless sensor networks," *Wireless Commun. Mobile Comput.*, vol. 2022, pp. 1–53, Mar. 2022.

[45] R. K. Yadav and R. P. Mahapatra, "Energy aware optimized clustering for hierarchical routing in wireless sensor network," *Comput. Sci. Rev.*, vol. 41, Aug. 2021, Art. no. 100417.

[46] P. S. Mehra, M. N. Doja, and B. Alam, "Fuzzy based enhanced cluster head selection (FBECS) for WSN," *J. King Saud Univ.-Sci.*, vol. 32, no. 1, pp. 390–401, 2020.

[47] P. S. Rathore, J. M. Chatterjee, A. Kumar, and R. Sujatha, "Energy-efficient cluster head selection through relay approach for WSN," *J. Supercomput.*, vol. 77, no. 7, pp. 7649–7675, Jul. 2021.

[48] A. García-Nájera, S. Zapotecas-Martínez, and K. Miranda, "Analysis of the multi-objective cluster head selection problem in WSNs," *Appl. Soft Comput.*, vol. 112, Nov. 2021, Art. no. 107853.

[49] A. A. Baradaran and K. Navi, "HQCA-WSN: High-quality clustering algorithm and optimal cluster head selection using fuzzy logic in wireless sensor networks," *Fuzzy Sets Syst.*, vol. 389, pp. 114–144, Jun. 2020.

LAXMINARAYAN SAHOO received the M.Sc. degree from Vidyasagar University, India, and the Ph.D. degree from The University of Burdwan, India. He is currently an Associate Professor of computer and information science with Raiganj University, Raiganj, India. He has published a good number of articles in international and national journals of repute. He is the author of the books *Advanced Operations Research* (Asian Books, New Delhi), *Advanced Optimization and Operations Research* (Springer Nature, Singapore). He edited a book titled *Real Life Applications of Multiple Criteria Decision-Making Techniques in Fuzzy Domain* (Springer Nature) and wrote several chapters from reputed publishers like Springer, IGI Global, CRC Press, Walter de Gruyter, McGraw-Hill, and Elsevier. His research interests include wireless sensor networks, distributed computing, reliability optimization, genetic algorithms, particle swarm optimization, graph theory, fuzzy game theory, interval mathematics, soft computing, fuzzy decision-making, and operations research. He is a fellow of ISROSET. He has received an MHRD Fellowship from the Government of India and received the Prof. M. N. Gopalan Award for Best Ph.D. thesis in Operations Research from the Operational Research Society of India (ORSI). He is a reviewer of several international journals and an Academic Editor of International Journal *Mathematical Problems in Engineering* (Hindawi).

SHRI SUPRIYAN SEN received the B.Sc. degree in computer science from the University of North Bengal, in 2008, and the M.C.A. degree from Vidyasagar University, India, in 2012. He is currently pursuing the Ph.D. degree with Raiganj University, West Bengal, India. He is an Assistant Professor with the Department of Computer and Information Science, Raiganj University, under the Government of West Bengal.

KALISHANKAR TIWARY received the B.Sc. (Hons.), M.Sc. (Hons.), and Ph.D. degrees from the University of Kalyani, West Bengal, India, and the D.Sc. degree in mathematics from Raiganj University.

He has a brilliant career. He was a Research Associate of CSIR. He is currently a Professor with the Department of Mathematics, Raiganj University, and the Dean of the Faculty of Science and Management. He is also the Director of the Centre for Advanced Studies in Mathematics (CASM). He has supervised a number of students for Ph.D. studies at different universities in West Bengal. He has research collaborations with eminent research workers in India and abroad. He has credit for more than 150 publications in India and abroad. He is associated with many research societies. Apart from this, he has a keen interest in history and literature. He has many articles and a book on political affairs as well. He is associated with the Lions Club and St. John's Ambulance. He has worked in different organizations and conferences as a committee member. He has participated in many research conferences. His current research interests include topology, functional analysis, fixed point theory, graph theory, and also different areas of computer and information science and computer science. He is a fellow of IARA, Trichy, India, and ISROSET, India. He has received a Prestigious Fellow Status from The Institute of Mathematics and Its application, U.K., and is referred to as a FIMA. He is an editorial board member in different journals of repute.

SARBAST MOSLEM is currently a Postdoctoral Research Fellow with University College Dublin, Ireland. He is also a highly accomplished researcher and educator in the field of decision-making. His extensive publication record includes numerous high-impact articles in leading international journals. He is a Principal Investigator of the VOTE-TRA project, funded by the Science Foundation of Ireland (SFI) and a National Project in Ireland (he obtained the grant, in 2023). He is involved in several Horizon 2020 projects, such as MoveCit and SENATOR.

TAPAN SENAPATI received the B.Sc., M.Sc., and Ph.D. degrees in mathematics from Vidyasagar University, India, in 2006, 2008, and 2013, respectively. He is currently a Postdoctoral Fellow with the School of Mathematics and Statistics, Southwest University, Chongqing, China. He has published three books and more than 120 articles in reputed international journals. His research results have been published in *Information Sciences*, *Fuzzy Sets and Systems*, IEEE TRANSACTIONS ON FUZZY SYSTEMS, IEEE TRANSACTIONS ON ENGINEERING MANAGEMENT, *Expert Systems with Applications*, *Applied Soft Computing*, *Engineering Applications of Artificial Intelligence*, *International Journal of Intelligent Systems*, and *International Journal of General Systems*. His main scientific interests concentrate on fuzzy sets, fuzzy optimization, soft computing, multi-attribute decision-making, and aggregation operators. He is a reviewer of several international journals and is also an Academic Editor of *Discrete Dynamics in Nature and Society* (SCIE), *Computational Intelligence and Neuroscience* (Scopus, Q1), and *Mathematical Problems in Engineering* (Scopus). He is also the Editor of the book *Real Life Applications of Multiple Criteria Decision-Making Techniques in Fuzzy Domain* (Springer). Recently, his name has been enlisted in the World's Top 2% Scientists list (Stanford University), from 2020 to 2021.