International Journal of Computer Networks and Applications (IJCNA)
DOI: 10.22247/ijcna/2025/48 Volume 12, Issue 5, September – October (2025)

RESEARCH ARTICLE

An Intelligent Neutrosophic Clustering based Optimized Routing Protocol for Wireless Sensor Networks

C.A. Kandasamy
Department of Computer Science, K.P.R College of Arts Science and Research, Coimbatore, Tamil Nadu, India.
kandaphd86@gmail.com

D. Maheshwari
Department of Computer Technology, KPR College of Arts Science and Research, Coimbatore, Tamil Nadu, India.
maheshwari.d@kprcas.ac.in

Received: 10 July 2025 / Revised: 24 September 2025 / Accepted: 02 October 2025 / Published: 30 October 2025

Abstract – Wireless Sensor Networks (WSNs) are a key technical innovation that might help accelerate the industrial revolution. In these networks, energy is the most important resource since sensor nodes are battery-powered, which cannot be replaced or recharged. As a result, the Cluster-based Routing Protocol (CRP) is regarded as one of the most successful methods for conserving energy and maximizing longevity in WSNs. However, increasing network lifespan in WSNs is heavily reliant on picking the right Cluster Heads (CHs), which may be difficult for decision makers. To overcome this problem, soft intelligence clustering approaches such as Fuzzy C-Means (FCM) and its variations were used. In these approaches, each node may be allocated a different membership degree relative to various clusters depending on network factors such as energy, distance, node density, and so on. However, these methods have a major weakness in that they equally weigh the contribution of network factors for cluster creation and CH selection, which is certainly influenced by irrelevant factors. This article may address this problem by employing the Neutrosophic C-Means (NCM) clustering approach. The benefit of NCM over FCM is that, although the fuzzy partition created by FCM can often only represent the uncertainty of clustering via membership, the neutrosophic partition produced by NCM can clearly disclose the imprecision and uncertainty of clustering decisions. However, NCM still weighs the relevance of each network parameter to clusters equally to pick the best CH. Hence, this study introduces a Self-adaptive Weighted Neutrosophic C-Means (SWNCM) clustering technique to cluster the network nodes and choose the best CH depending upon different factors, with mean distance from the Base Station (BS) and cluster member nodes, residual energy, node density, node degree, delay, etc. Additionally, a new Greylag Goose Optimization (GGO) technique is implemented to choose the ideal path from CHs to the BS according to the above-mentioned parameters. This entire proposed work is named as an Intelligent Neutrosophic Clustering-based Routing Protocol (INCRP). Consequently, this INCRP supports improving energy use and extending the network lifespan by handling uncertainty in WSNs. Finally, the findings demonstrate that the suggested INCRP surpasses other conventional clustering techniques for different network metrics.

Index Terms – WSN, CH Selection, DBSCAN, K-Means, Neutrosophic C-Means, Greylag Goose Optimization.

1. INTRODUCTION

In a Wireless Sensor Network (WSN), numerous small, inexpensive sensor nodes are dispersed across a predetermined area at random [1]. Target monitoring, disaster recovery, and space exploration are just a few of the many WSN uses [2]. However, charging or replacing the nodes can be extremely expensive or even impossible in hostile circumstances [3]. Consequently, energy conservation for battery-powered nodes has always been crucial for WSNs to maximize their operational longevity. The CRPs have been rigorously tested as an effective approach for minimizing energy consumption, augmenting scalability, lowering latency, and prolonging network longevity [4]. Dynamically dividing the network's nodes into a predetermined number of clusters is how clustering is applied in CRPs [5]. Cluster management, data collecting, aggregation, and transmission are all responsibilities of the CH, a node with optimal resources in each cluster. The other nodes, known as Cluster Members (CMs), provide their sensed data to the CH [6]. Additionally, all CH data is collected and processed by the Base Station (BS). In this light, the role of the CHs as a go-between for the CMs and the BS becomes even clearer. For this reason, CH selection is considered an NP-hard problem, despite its centrality to CRPs [7]. This problem has led to the increased use of various metaheuristic algorithms, including Genetic Algorithm (GA), Particle Swarm Optimization (PSO), Artificial Bee Colony (ABC), etc. [8, 9]. These optimization strategies have improved data delivery

RESEARCH ARTICLE

performance, reduced energy consumption, and enhanced network efficiency as compared to Low Energy Adaptive Clustering Hierarchy (LEACH), the most prevalent CRP in WSNs. Still, these algorithms only consider residual energy when selecting CHs. Using such a narrow criterion for selection causes cluster formation disruptions frequently, which in turn depletes important resources. Thus, hierarchical clustering algorithms, K-means, and K-medoids have been employed to select the optimal CHs according to various network criteria, such as distance, node density, cluster stability, residual energy, and so on [10].

On the other hand, they are vulnerable to data noise and outliers. Network parameters in WSNs often have fuzzy values, which means they are imprecise and unpredictable. As a result, Fuzzy Logic (FL) systems are often needed for accurate CH elections [11, 12]. Integrating uncertainty considerations for WSNs is vital in designing resilient, adaptable, and dependable CRPs that can function efficiently in dynamic and unpredictable settings. In most cases, a range of possible values is used to depict the uncertainty in network parameters, with the most likely or average value in the middle.

Such values are defined using various fuzzy membership functions in FL systems, such as Triangular Fuzzy Numbers (TFNs), trapezoidal, Gaussian fuzzy numbers, and extended bell-shaped fuzzy numbers. TFNs are best suited for discrete values, whereas other fuzzy number representations are employed for continuous parameters. According to this context, Sahoo et al. [13] suggested the CRP by integrating clustering methods with Multi-Criteria Decision Making (MCDM) strategies. To begin, the optimal number of clusters was determined using the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) approach, which employed the elbow scheme and K-means clustering. After that, the best CHs for each cluster were selected using the MCDM method, which prioritized cluster creation.

Furthermore, they employed TFNs to manage ambiguity and imprecision in the criteria for CH selection according to the network metrics, including distance from BS, mean distance of cluster nodes, dependability, and remaining energy. However, TFNs aren't up to the task of dealing with complicated uncertainties, especially in massive WSNs. Adopting FCM clustering schemes can resolve this limitation by assigning nodes to multiple clusters with varying membership degree values. In these methods, each node may be assigned a distinct membership degree relative to multiple clusters based on network characteristics such as energy, distance, node density, and so on. Nevertheless, these methods have a significant drawback in that they equally consider the contribution of network characteristics to cluster creation and CH selection, which is undoubtedly influenced by irrelevant factors.

This article may solve this issue by using the NCM clustering technique. The advantage of NCM over FCM is that, whereas the fuzzy partition established by FCM can frequently just convey the uncertainty of clustering via membership, the neutrosophic partition produced by NCM can clearly expose the imprecision and uncertainty of clustering decisions. In NCM, the neutrosophic membership function provides clusters with support for each network parameter to reflect uncertainty. Meanwhile, imprecision is represented by the imprecise cluster (a disjunction of multiple distinct clusters).

This approach effectively represents the uncertainty and imprecision in clustering findings induced by cluster overlap. To choose the optimal CH, NCM rates the significance of each network parameter to clusters equally. As a result, this research presents the SWNCM clustering approach, which is used to cluster network nodes and choose the optimal CH based on many variables. Furthermore, the GGO algorithm is used to choose the optimal route from the CHs to the BS, resulting in great energy efficiency and network lifespan.

1.1. Main Contributions

This manuscript proposes the INCRP for extending the lifetime of WSN. This protocol involves two major phases: (i) clustering and CH selection, and (ii) data transmission. The key contributions of these phases include:

1. In clustering and CH selection phase, the network nodes are clustered using the SWNCM clustering algorithm based on the different factors, including distance from BS, mean distance of cluster nodes, reliability, remaining energy, node density, delay, etc. Once the cluster is formed, the best node using the fore-mentioned factors in individual cluster is chosen as the optimal CH.
2. During data transmission, the GGO algorithm chooses the best route from CHs to the BS by taking into account the aforementioned characteristics to increase network endurance and energy effectiveness.

In this order, the following sections are organized: A review of the relevant literature is covered in Section 2. Chapter 3 explains the INCRP for WSNs, while Chapter 4 displays the simulation results. The research is ended and recommendations for future improvements are given in Section 5.

2. LITERATURE SURVEY

2.1. Clustering and Routing Based on Machine Learning Algorithms

To address uncertainty in WSN, Sen et al. [14] created a novel algorithm that combines the MCDM technique with K-medoids clustering. Using the K-medoids approach, the whole network was divided into several clusters. The MCDM approach, specifically the TOPSIS method, was used to select

CHs from each cluster based on variables like distance from the sink, average distance of the cluster nodes, cluster dependability, and residual energy. This strategy led to a longer lifespan for the network. Despite the multi-dimensional node placements, the 2D node location was considered, which might affect the energy used for data transport throughout the network and the longevity of the whole network in real-time situations.

A Multi-Layer Perceptron (MLP) neural network model for effective CH selection in WSNs was created by Jalili et al. [15]. There were two parts to this model: training and testing. To determine the optimal CH position based on the sensor's circumstances, the trained model was used during testing. This model was trained using the necessary parameters during training. Using clustering and CH selection, this approach drastically enhances WSN performance by preventing wasteful transmissions and receptions. Nevertheless, node energy consumption remained high.

2.2. Clustering and Routing Based on Metaheuristic Algorithms

The Genetic Algorithm (GA) was proposed by Pravin et al. [16] to reduce energy consumption during network operations. Heterogeneous WSNs used the Stochastic CH Selection Model (SCHSM), which takes into account distance, node density, energy, and node capacity to determine the GA's fitness function. To minimize transmission delays between sensors and sinks, this model was created for several mobile nodes that serve as sinks. However, the cluster's lifespan was limited.

Using optimization methods, Divya and Sudhakar [17] attempted to determine the optimal path and choose the most suitable CH for the WSN. To choose CHs and non-CHs, the Adaptive Multi-Path Routing Protocol (AMPRP) using the Circle-Inspired Optimization Algorithm (CIOA) was introduced. By calculating the fitness function with respect to energy, time, distance, and traffic density, the CIOA was able to choose between the two CHs with the best possible probability. The same CIOA was also used to determine the shortest route based on the results. While the number of nodes increased, throughput remained poor, and time complexity remained high.

To choose the best CH for WSNs, Alshammri [18] created a novel method based on the Improved Squirrel Search Algorithm based Clustering (ISSA-C). A number of improvements were made to the standard SSA with the goal of increasing solution quality and speeding up convergence. The exploration and exploitation capabilities of the SSA were enhanced by executing a local search algorithm, dynamic step size control, and adaptive population initialization. Factors such as intra-cluster distance, sink distance, residual energy, and the average CH balance constituted the basis of the fitness function for this method. But there were a lot of power consumption and a short lifespan for the network.

2.3. Clustering and Routing Based on Hybrid Algorithms

In order to decrease and balance energy usage in WSNs, Wang [19] created Distributed PSO a Fuzzy Clustering Protocol (DPFCP). This DPFCP's CHs were identified using the Mamdani FL approach, which takes into account the BS distance, middle point, leftover energy, and node degree. The PSO method was used to optimize the fuzzy criteria in order to select the best nodes to serve as CHs. Based on factors such as distance to the CH, leftover energy, and variation in the CM of the CH, non-CH joining clusters were established once the CHs were selected. In addition, an on-demand technique was used to reduce energy usage by maintaining clusters locally and globally using local information. On the other hand, the latency and network scalability are both negatively impacted by CHs communicating with the BS in single-hop mode.

To optimize the fuzzy rules, Yuebo et al. [20] developed an improved PSO approach and used it to construct a Novel Fuzzy CRP (NFCRP). After considering distance to centrality, node degree deviation, and residual energy, a Fuzzy Inference System (FIS) was used to ascertain the optimal grouping. This may reduce energy consumption inside clusters and produce optimal clusters. Reduced energy consumption between clusters was the consequence of switching to a different FIS that prioritized energy balance, distance to the BS, and data load variance while deciding on the best route. In addition, the upgraded PSO, with parameters updated by the chaotic mapping and adaptive inertia weight, altered the rules of both FISs. Unfortunately, not taking node mobility into account affected network resilience and dependability.

Manoharan et al. [21] enhanced WSN with adaptive Entropy-Bald Eagle Search (EBES) and Density-based Adaptive Soft (DAS), resulting in an ideal low-power routing system. At first, the DAS clustering approach was used to group the nodes. The next stage involved selecting the CH using an updated beetle swarm optimization algorithm that considers distance, energy, trust, and throughput. The values of the entropy weights of the nodes were calculated, and the node with the highest entropy was selected to transmit the data. In addition, an adaptive mechanism was used in EBES optimization to ensure optimum data transmission. Nevertheless, mobility factors were not considered, which impacts the network energy efficiency significantly.

Jing [22] introduced a Harris Hawks Optimization Clustering with Fuzzy Routing (HHOCFR). An improved HHO algorithm was used to choose the optimum CHs and simultaneously form perfect clusters using an encoding technique. Furthermore, the FL was employed to determine

RESEARCH ARTICLE

the relay CH by considering the residual energy, fluctuations in the cluster's energy usage, and the proximity to the crossover point. Furthermore, the rotation of CHs within clusters and the initiation of re-clustering in the network were executed using an adaptive cluster management technique to avert repetitive re-clustering. However, other aspects, such as node density and traffic factors, were not addressed, which leads to a degradation in lifetime and energy efficiency. Using K-means clustering and the firefly algorithm, Chen [23] proposed a new method for optimizing the WSN topology. Nodes were partitioned using the K-means clustering approach, which minimizes the within-cluster variance. The search was guided using the firefly algorithm, an optimization tool based on swarm intelligence that mimics the flashing interaction between fireflies. The nodes were adjusted dynamically according to traffic circumstances using the firefly algorithm. However, network energy use was high.

Chai et al. [24] introduced the Energy-efficient Scalable Routing that cluster sensor node using Hierarchical Agglomerative Clustering (ESR-HAC) to improve network lifetime and energy efficiency among sensor nodes. This technique utilized a hierarchical clustering algorithm to reduce transmission cost.

A cost function was formed to lessen communication overhead between member nodes and CHs, considering the coverage range, connection, and remaining energy of the sensor nodes. The introduction of GA allowed for optimal inter-cluster routing, balancing energy usage, and eliminating hot spots.

However, for large-scale networks, Packet Delivery Rate (PDR) and robustness were poor. Table 1 summarizes the above-studied algorithms according to their benefits, drawbacks, and performance measures.

Table 1 An Overview of Existing WSN Routing and Clustering Methods

Ref. No.	Algorithms	Benefits	Drawbacks	Performance measures
[14]	MCDM with K-medoids clustering	It achieved better energy efficiency.	Energy consumption was high during data transfer.	No. of rounds = 2000: No. of nodes alive = 20
[15]	MLP	It reduced packet loss efficiently.	Energy consumption of nodes remained high.	Time = 300 sec: No. of packet loss = 8%; Energy consumed = 100J
[16]	SCHSM using GA	It increased PDR and reduced transmission delay significantly.	Cluster lifetime was low.	Node density = 150: Packet Delivery Ratio (PDR) = 410000 kilobits per second (kbps); Overhead = 0.4; Cluster lifetime = 0.5 milliseconds (ms); Transmission delay = 3 ms; Residual energy = 9.86 J
[17]	AMPRP-CIOA	It attained better throughput and energy consumption.	The time complexity was high while increasing the node density.	No. of nodes = 1000: Energy consumption = 0.2; Network lifetime = 3.24×10^5 Throughput = 0.92
[18]	ISSA-C	It reduced end-to-end delay significantly.	Energy consumption was high, and network lifetime was low.	No. of nodes = 1000: PDR = 88%; Throughput = 0.5 bits per second (bps);

				Energy usage = 210 J; End-to-end delay = 10 ms
[19]	DPFCP	It can increase the network lifetime and throughput.	Network energy consumption remained high.	No. of nodes = 100: Network throughput = 8x10^8 No. of rounds = 2000: Alive nodes = 5
[20]	NFCRP	Its network throughput and longevity were both maximized.	The failure to take the node's mobility aspects into account resulted in persistently high energy usage.	No. of rounds = 100: Alive nodes = 100 No. of nodes = 100: Network throughput = 14 x10^7; Energy consumption = 12 Joules (J)
[21]	EBES and DAS	It ensured optimum data transfer with higher throughput.	Mobility factors were not considered, which impacts the network energy efficiency significantly.	Delay = 6.5 ms; Throughput = 320.1 kbps; Energy consumption = 1.92 J
[22]	HHOCFR	It has a high network throughput.	Other factors like node density and traffic density were not considered, which leads to a degradation of the network lifetime and energy use.	No. of nodes = 100: Network throughput = 4.45x10^8 No. of rounds = 1600: Energy usage = 100 J
[23]	K-means clustering and firefly algorithm	It can attain better throughput by transmitting a higher quantity of data packets.	Network energy consumption was high.	Network energy consumption = 65 J; Residual energy = 20 J; No. of data packets = 90
[24]	ESR-HAC	It can balance the network energy consumption and data delivery.	PDR and robustness were limited for large complex networks.	No. of nodes = 400: Packets received by the sink = 3.7x10^6; Packet delivery = 95%

2.1. Research Gap

From this literature survey, it can be inferred that existing CRPs focus on enhancing energy efficacy and network lifespan using different approaches, including machine learning, metaheuristic, or hybrid. However, most of these protocols still cannot handle uncertainty in node parameters, which impacts the decision about appropriate CH selection and overall network performance. Hence, this study addresses the uncertainty problems by developing a new INCRP that considers various factors to enhance both CH and route selection in large-scale WSNs. The main differences between the existing and the proposed INCRP are related to how the uncertainties are treated, factor weighting, and the extent of optimization. Unlike previous models, such as DPFCP, HHOCFR, etc., the proposed INCRP incorporates the SWNCM clustering, which characterizes the membership of neutrosophic sets to represent truth, indeterminacy, and falsity. This method allows factors such as residual energy, node density, delay, and traffic density to be weighted differently, unlike the equal treatment of factors in K-medoid or K-means, the reduction of interference by irrelevant factors, and increased accuracy of clustering in uncertain settings. Further, INCRP employs GGO to select the route by balancing exploration and exploitation in a multi-factor fitness function compared to the single-objective optimizations like I-SSA and ESR-HAC. Therefore, the

RESEARCH ARTICLE

INCRP achieves higher energy efficiency, throughput, and network lifetime.

3. PROPOSED METHODOLOGY

This part explains the suggested INCRP algorithm. Figure 1 presents the schematic pipeline of this study.

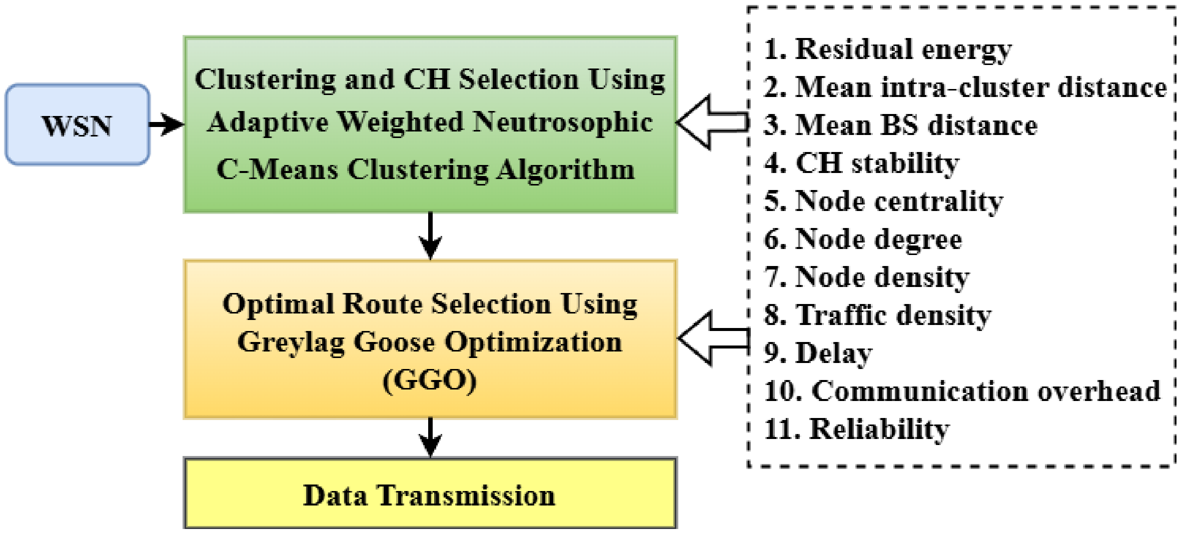

Figure 1 Schematic Representation of the Proposed Study

3.1. Network Model

The WSN is composed of n sensor nodes and a single BS. There are two categories into which every node falls: (i) regular nodes and (ii) CH nodes. The normal nodes observe environmental information and convey the sensing data to CHs. Th CH node is selected from these normal nodes. The CH receives, aggregates, and transfers the information to the BS. Below hypotheses are taken into account in this study:

- Nodes are placed randomly and in an equitable distribution inside a square.
- BS nodes form the backbone of the hierarchical WSN. Here, the user and the sensor network are able to communicate with one another. Communication with nodes exposed to multipath attenuation is feasible since the BS is outside the square. Node-to-node communication is unaffected by multi-path debilitation.
- Due to the similar capacities and starting battery levels, the nodes work together harmoniously, modifying their operations according to the time of day.
- A node may communicate with any other node, including BS.
- All nodes stay motionless or immovable.
- Every node sends a signal with a consistent duration while observing its surroundings.
- WSN parameters have uncertainty and are defined by fuzzy values.

3.2. Energy Model

The energy model is commonly referred to as the first order radio concept. This study only concentrates on energy use when the data is communicating. The total energy usage contains the energy transmission resulting from information sensing, aggregation, and allocation.

In this model, a normal node and CH node exchange L-bit data at distance d, as well as the consumption of energy is determined using equations (1) and (2).

$$E_{TX}(L,d) = E_{elec} \times L + \varepsilon_{fs} \times L \tag{1}$$

$$E_{RX}(L) = E_{elec} \times L \tag{2}$$

In equations (1) and (2), $E_{TX}(L,d)$ is the energy spent when exchanging information of size L bits, and $E_{RX}(L)$ is the energy spent when receiving the data. The energy utilization of the amplifier during data transfer is calculated by equations (3) and (4).

$$\varepsilon_{amp} = \begin{cases} \varepsilon_{fs} d^2, & d \le d_0 \\ \varepsilon_{mp} d^4, & d > d_0 \end{cases} \tag{3}$$

$$\text{Where } d_0 = \sqrt{\frac{\varepsilon_{fs}}{\varepsilon_{mp}}} \tag{4}$$

The threshold distance, denoted as d_0, and the free-space as well as multi-path amplification constants, ε_{fs} and ε_{mp}, respectively, are used in equations (3) and (4). If $d > d_0$, the normal node can utilize multipath fading channel; otherwise, sensor node can utilize free-space propagation model.

RESEARCH ARTICLE

3.3. Formulation of Fitness Function

The fitness function plays a significant role in optimizing routes, selecting CHs and determining the efficacy of network operation in WSNs. The fitness function is composed of various elements, each of which contributes to a distinct aspect of network and energy effectiveness. These variables are sensibly selected to satisfy detailed problems about energy consumption, distance to BS, and cluster size balance.

3.3.1. Residual Energy

As the network's lifetime is determined by the energy utilized, reducing energy usage is essential. The value is calculated by adding the current energy of all selected CHs. Because the total amount of energy is essentially maximized, considering the reciprocal assures that all objective functions are perfectly balanced. The proposal algorithm aims to identify CHs that enhance overall network energy efficiency by incorporating the remaining energy of each CH into the fitness function. The reciprocal of the total energy of all selected CHs is computed as f_1 in equation (5).

$$f_1 = \frac{1}{\sum_{j=1}^{m}\left(E_{CH_j}\right)} \tag{5}$$

In equation (5), m is the CHs count and E_{CH_j} is the current energy of j^{th} CH.

3.3.2. Mean Intra-Cluster Distance

A node's intra-cluster distance is the sum of its distances from all other nodes in the sensor cluster. In order for the network to use as few resources as possible, it is crucial to reduce the intra-cluster distance. In order to reduce the distance between sensor nodes and their linked CHs inside the same cluster, the mean intra-cluster distance factor f_2 is determined using equation (6).

$$f_2 = \sum_{j=1}^{m}\left(\frac{1}{l_j} D_{sn-CH_j}\right) \tag{6}$$

The distance among sensor node (sn) as well as the j^{th} CH is denoted as D_{sn-CH_j} in equation (6), while l_j represents the count of sensor nodes in the j^{th} cluster.

3.3.3. Mean BS Distance

Divide the total number of sensor nodes linked to the CH by the distance between the CH and the BS to get the mean BS distance. Because of its significance, the distance between CHs and BS is a major component in energy consumption, and the BS distance factor (f_3) takes this into account. It is determined by equation (7).

$$f_3 = \sum_{j=1}^{m}\left(\frac{1}{l_j} D_{CH_j-BS}\right) \tag{7}$$

In equation (7), D_{CH_j-BS} represents the distance among j^{th} CH as well as BS.

3.3.4. CH Stability

Maintaining a balance in cluster size is critical since a higher node count inside the cluster results in increased energy dissipation. The positioning of sensor nodes enables the formation of clusters of varied sizes, both small and large, to optimize energy usage. The CH stability f_4 in equation (8) underlines the need of having equal-sized clusters in the network.

$$f_4 = \sum_{j=1}^{m} \frac{n}{m} - l_j \tag{8}$$

Equation (8) uses the total number of sensor nodes in the network, denoted as n.

3.3.5. Node Centrality

The centrality of node (sn_i) is determined by equation (9).

$$f_5 = \frac{\sum_{i,k=1}^{n}\left(D_{sn_i-sn_k}\right)}{|Neighbor(sn_i)|}, k \neq i \tag{9}$$

In equation (9), $|Neighbor(sn_i)|$ is the collection of neighbors for node sn_i inside its communication range R and $D_{sn_i-sn_k}$ represents the distance among node sn_i and its adjacent node sn_k.

3.3.6. Node Density

Nodes in the CH region are more densely packed, which aids in energy conservation and protection. The nodes' energy is regularly utilized, which is primarily due to intra-cluster transmission. As a result, a trustworthy CH node is chosen from the current nodes. Node density f_6, also known as the average distance between neighboring nodes within a cluster, is computed using equation (10).

$$f_6 = \frac{\sum_{i=1}^{n} D_{sn_i-N_N}(i)}{n} \times \frac{1}{D_{CH-BS}} \tag{10}$$

In equation (10), $D_{sn_i-N_N}(i)$ refers to the distance among i^{th} node and its neighbor nodes (N_N), and D_{CH-BS} refers to the Euclidean distance among the CH as well as BS.

3.3.7. Node Degree

It is computed using equation (11).

$$f_7 = \frac{\sum_{i=1}^{C} \sum_{j\in NeighborCH(CH_j)} |ClusterSet(CH_j)|}{\sum_{i=1}^{C} \Psi(CH_i)} \tag{11}$$

In equation (11), C is the cluster count, $\Psi(CH_i)$ is the hop count from CH_i to the BS, $|ClusterSet(CH_j)|$ is the number of its members, and $NeighborCH(CH_j)$ is the set of neighbor CHs of CH_j.

3.3.8. Traffic Density (TD)

A certain degree of TD is essential to run an effective network. The key factors influencing TD are channel load, buffer use, and packet loss. The TD of node (sn_i), f_8 is calculated by averaging these three factors, as shown in equation (12).

$$f_8 = \sum_{i=1}^{n}\left(\frac{1}{3}[BU_i + Q_i + PD_i]\right) \quad (12)$$

$$\text{Where } BU_i = \frac{B_{space}}{B_{size}} \quad (13)$$

$$Q_i = \frac{Q_l}{Q_y} \quad (14)$$

$$PD_i = \frac{D_{busy}}{S} \quad (15)$$

In equations (12)-(15), BU_i is the buffer use of i^{th} sensor node, B_{space} is the available buffer space, B_{size} is the total buffer size, Q_i is the Q_l is the queue length, Q_y is the maximum queue size, S is the total rounds executed during the simulation, D_{busy} is the number of busy channels.

3.3.9. Delay

To choose efficient CH, it is vital to minimize network delay, which is strongly linked to the total nodes within a certain cluster. The optimal cluster volume can be achieved by limiting the total number of nodes in every cluster, as the delay increases as the number of cluster participant's increases. So, the CH has fewer members to reduce network delay. The delay's fitness f_9 is calculated in equation (16).

$$f_9 = \frac{\max\left(\left\|DL_{sn_i} - DL_{sn_j}\right\| \times l_j\right)}{l_j} \quad (16)$$

In equation (16), DL_{sn_i} and DL_{sn_j} are the delay of i^{th} and j^{th} sensor nodes.

3.3.10. Communication Overhead

It is impacted by the packet size and distance. This is determined by the packet size to hops needed for delivery ratio, as given in equation (17.)

$$f_{10} = \sum_{i=1}^{n}\frac{S \times H}{DR} \quad (17)$$

In equation (17), S denotes the packet size, H is the hop count needed for transmission, and DR is the PDR.

3.3.11. Reliability

It is an average measure across all clusters and calculated in equation (18).

$$f_{11} = \frac{1}{C}\sum_{j=1}^{C}\left(1 - \frac{\text{ClusterSet}(CH_j)}{\text{ClusterSet}(CH_j)_{max} + 1}\right) \quad (18)$$

In equation (18), $\text{ClusterSet}(CH_j)_{max}$ denotes the maximum nodes in cluster j.

Therefore, a fitness function for CH selection and path optimization is defined in equation (19).

$$F = \sum_{i=1}^{11} \alpha_i f_i \quad (19)$$

In equation (19), α_i is the weight coefficient for i^{th} factor in the fitness function.

3.4. Self-Adaptive Weighted Neutrosophic C-Means Clustering Algorithm

Consider a WSN with a series of sensor nodes $n = \{x_1, \dots, x_n\}$, characterized by various factors $A = \{a_1, \dots, a_S\}$. These nodes are grouped into C clusters within a framework of discernment $\theta = \{\theta_1, \dots, \theta_C\}$. In this study, to better differentiate the significance of each factor defined in fitness function, a novel SWNCM clustering protocol is introduced. This algorithm assumes that numerous elements, like remaining energy, node degree, centrality, reliability, etc., have distinct contributions to the CH selection within each cluster. This results in reducing the negative impact of less significant factors on the clustering and CH selection processes. Besides, it handles uncertainty and imprecision in clustering outcomes, especially when clusters are overlapped.

The objective function of this SWNCM is designed to optimize clustering and CH selection when minimizing uncertainty. It is defined in equations (20) to (22).

$$\mathcal{J}(T, I, F, V, W) = \sum_{i=1}^{n}\sum_{j=1}^{C}\sum_{s=1}^{S}\{\varpi_1 T_{ij}\}^{\beta}\omega_{j,s}\left\|x_{i,s} - v_{j,s}\right\|^2 + \sum_{i=1}^{n}\sum_{s=1}^{S}\{\varpi_2 I_i\}^{\beta}\omega_{imax,s}\left\|x_{i,s} - v_{imax,s}\right\|^2 + \delta^2\sum_{i=1}^{n}\{\varpi_3 F_i\}^{\beta} + \gamma\sum_{s=1}^{S}\left(\sum_{j=1}^{C}\omega_{j,s}\log\omega_{j,s} + \omega_{imax,s}\log\omega_{imax,s}\right) \quad (20)$$

$$\text{s.t. } \sum_{j=1}^{C} T_{ij} + I_i + F_i = 1, 0 < T_{ij}, I_i, F_i < 1;\ \sum_{s=1}^{S}\omega_{j,s} = 1;\ \sum_{s=1}^{S}\omega_{imax,s} = 1$$

$$\text{with } v_{imax,s} = \frac{v_{p_i,s} + v_{q_i,s}}{2} \quad (21)$$

$$p_i = \underset{j=1,\dots,C}{\operatorname{argmax}}(T_{ij}),\ q_i = \underset{j \neq p_i \cap j=1,\dots,C}{\operatorname{argmax}}(T_{ij}) \quad (22)$$

From equation (20) to (22), T, I, F define the neutrosophic partition matrix components namely membership of truth, membership of indeterminacy, as well as the membership of falsity functions, respectively), V is the cluster center matrix, W is the feature (i.e., network factors) weight matrix, $\varpi_1, \varpi_2, \varpi_3$ are the weight parameters, β is the weighting exponent, δ is the noise threshold, γ is the variable of feature weight, $\|\cdot\|$ is the distance measure, $\omega_{j,s}$ and $\omega_{imax,s}$ are the contributions of s -dimensional factors to specific and imprecise clusters, respectively. γ is a positive regularization variable, T_{ij}, I_i, and F_i are the support degree of x_i to specific (precise) cluster θ_j, boundary (imprecise) cluster θ_{imax}, and outlier, respectively. $x_{i,s}, v_{j,s}$, and $v_{imax,s}$ are the s^{th} feature of x_i, v_j, and v_{imax}, respectively, where v_j and v_{imax} are the j^{th} cluster center and imprecise cluster center for x_i, respectively.

A Lagrange optimization technique, such as Lagrange multiplier λ_i, is employed to minimize equation (20), which involves updating one variable while keeping the others constant.

3.4.1. Network Partition Update

The objective function in equation (20) can be redefined as:

$$\mathcal{L}(T, I, F, \lambda_i) = \mathcal{J} - \sum_{i=1}^{n} \lambda_i \left(\sum_{j=1}^{C} T_{ij} + I_i + F_i - 1\right) \quad (23)$$

Calculating the derivatives of equation (23) w.r.t. T_{ij}, I_i, F_i, and λ_i, respectively, and assigning them to zeros. As a result, equations (24) to (27) are obtained.

$$\frac{\partial \mathcal{L}}{\partial T_{ij}} = \sum_{s=1}^{S} \beta \{\varpi_1 T_{ij}\}^{\beta-1} \omega_{j,s} \|x_{i,s} - v_{j,s}\|^2 - \lambda_i = 0 \quad (24)$$

$$\frac{\partial \mathcal{L}}{\partial I_i} = \sum_{s=1}^{S} \{\varpi_2 I_i\}^{\beta-1} \omega_{imax,s} \|x_{i,s} - v_{imax,s}\|^2 - \lambda_i = 0 \quad (25)$$

$$\frac{\partial \mathcal{L}}{\partial F_i} = \beta \delta^2 \{\varpi_3 F_i\}^{\beta-1} - \lambda_i = 0 \quad (26)$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_i} = \sum_{j=1}^{C} T_{ij} + I_i + F_i - 1 = 0 \quad (27)$$

From equations (24) to (27),

$$T_{ij} = \sum_{s=1}^{S} \frac{1}{\varpi_1} \left(\frac{\lambda_i}{\beta}\right)^{\frac{1}{\beta-1}} \left(\frac{1}{\omega_{j,s}\|x_{i,s} - v_{j,s}\|^2}\right)^{\frac{1}{\beta-1}} \quad (28)$$

$$I_i = \sum_{s=1}^{S} \frac{1}{\varpi_2} \left(\frac{\lambda_i}{\beta}\right)^{\frac{1}{\beta-1}} \left(\frac{1}{\omega_{imax,s}\|x_{i,s} - v_{imax,s}\|^2}\right)^{\frac{1}{\beta-1}} \quad (29)$$

$$F_i = \frac{1}{\varpi_3} \left(\frac{\lambda_i}{\beta}\right)^{\frac{1}{\beta-1}} \left(\frac{1}{\delta^2}\right)^{\frac{1}{\beta-1}} \quad (30)$$

Moreover, equations (31) to (34) are obtained.

$$T_{ij} = \frac{1}{\varepsilon \varpi_1} \sum_{s=1}^{S} \left(\omega_{j,s} \|x_{i,s} - v_{j,s}\|^2\right)^{-\frac{1}{\beta-1}} \quad (31)$$

$$I_i = \frac{1}{\varepsilon \varpi_2} \sum_{s=1}^{S} \left(\omega_{imax,s} \|x_{i,s} - v_{imax,s}\|^2\right)^{-\frac{1}{\beta-1}} \quad (32)$$

$$F_i = \frac{1}{\varepsilon \varpi_3} \delta^{-\frac{2}{\beta-1}} \quad (33)$$

$$\varepsilon = \frac{1}{\varpi_1} \sum_{j=1}^{C} \sum_{s=1}^{S} \left(\omega_{j,s} \|x_{i,s} - v_{j,s}\|^2\right)^{-\frac{1}{\beta-1}} + \frac{1}{\varpi_2} \sum_{s=1}^{S} \left(\omega_{imax,s} \|x_{i,s} - v_{imax,s}\|^2\right)^{-\frac{1}{\beta-1}} + \frac{1}{\varpi_3} \delta^{-\frac{2}{\beta-1}} \quad (34)$$

3.4.2. Cluster Centers Update

An unconstrained optimization issue is presented by minimizing the value of equation (20) with respect to V. The following clustering models can be obtained by taking the derivative of equation (20) with respect to V and setting it equal to zero as shown in equation (35):

$$\frac{\partial \mathcal{L}}{\partial v_{j,s}} = \sum_{i=1}^{n} -2\{\varpi_1 T_{ij}\}^{\beta} \omega_{j,s}(x_{i,s} - v_{j,s}) = 0 \quad (35)$$

Hence, equation (36) is obtained.

$$v_{j,s} = \frac{\sum_{i=1}^{N} \{\varpi_1 T_{ij}\}^{\beta} x_{i,s}}{\sum_{i=1}^{N} \{\varpi_1 T_{ij}\}^{\beta}} \quad (36)$$

3.4.3. Feature Weights Update

In order to resolve the constrained minimization issue with respect to W, Lagrange multipliers ϱ_j as well as ρ are used. Given equation (20), the objective function has the potential to transform into equation (37),

$$\mathcal{L}(W, \varrho_j, \varrho_{imax}) = \mathcal{J} - \sum_{j=1}^{C} \varrho_j \left(\sum_{s=1}^{S} \omega_{j,s} - 1\right) - \rho\left(\sum_{s=1}^{S} \omega_{imax,s} - 1\right) \quad (37)$$

Taking derivatives of equation (37) w.r.t. $\omega_{j,s}, \omega_{imax,s}, \varrho_j$, and ρ, respectively, and assigning them to zeros. So, equations (38) to (41) is obtained.

$$\frac{\partial \mathcal{L}}{\partial \omega_{j,s}} = \sum_{i=1}^{n} \{\varpi_1 T_{ij}\}^{\beta} \|x_{i,s} - v_{j,s}\|^2 + \gamma(\ln \omega_{j,s} + 1) - \varrho_j = 0 \quad (38)$$

$$\frac{\partial \mathcal{L}}{\partial \omega_{imax,s}} = \sum_{i=1}^{n} \{\varpi_2 I_i\}^{\beta} \|x_{i,s} - v_{imax,s}\|^2 + \gamma(\ln \omega_{imax,s} + 1) - \rho = 0 \quad (39)$$

$$\frac{\partial \mathcal{L}}{\partial \varrho_j} = \sum_{s=1}^{S} \omega_{j,s} - 1 = 0 \quad (40)$$

$$\frac{\partial \mathcal{L}}{\partial \rho} = \sum_{s=1}^{S} \omega_{imax,s} - 1 = 0 \quad (41)$$

Moreover, equations (42) to (45) are obtained.

$$\omega_{j,s} = \frac{\exp(-\gamma^{-1}\mathcal{H}_{j,s})}{\sum_{t=1}^{S} \exp(-\gamma^{-1}\mathcal{H}_{j,t})} \quad (42)$$

$$\omega_{imax,s} = \frac{\exp(-\gamma^{-1}\mathcal{H}_{imax,s})}{\sum_{t=1}^{S} \exp(-\gamma^{-1}\mathcal{H}_{imax,t})} \quad (43)$$

$$\mathcal{H}_{j,s} = \sum_{i=1}^{n} \{\varpi_1 T_{ij}\}^{\beta} \|x_{i,s} - v_{j,s}\|^2 \quad (44)$$

$$\mathcal{H}_{imax,s} = \sum_{i=1}^{n} \{\varpi_2 I_i\}^{\beta} \|x_{i,s} - v_{imax,s}\|^2 \quad (45)$$

Thus, this SWNCM clustering algorithm can cluster the sensor nodes. Once the cluster formed, the ideal node in individual cluster depending on the fitness function is chosen as CH. An entire procedure is clearly described in Algorithm 1 and the flowchart of this algorithm is portrayed in Figure 2. Table 2 lists the parameters used in SWNCM.

RESEARCH ARTICLE

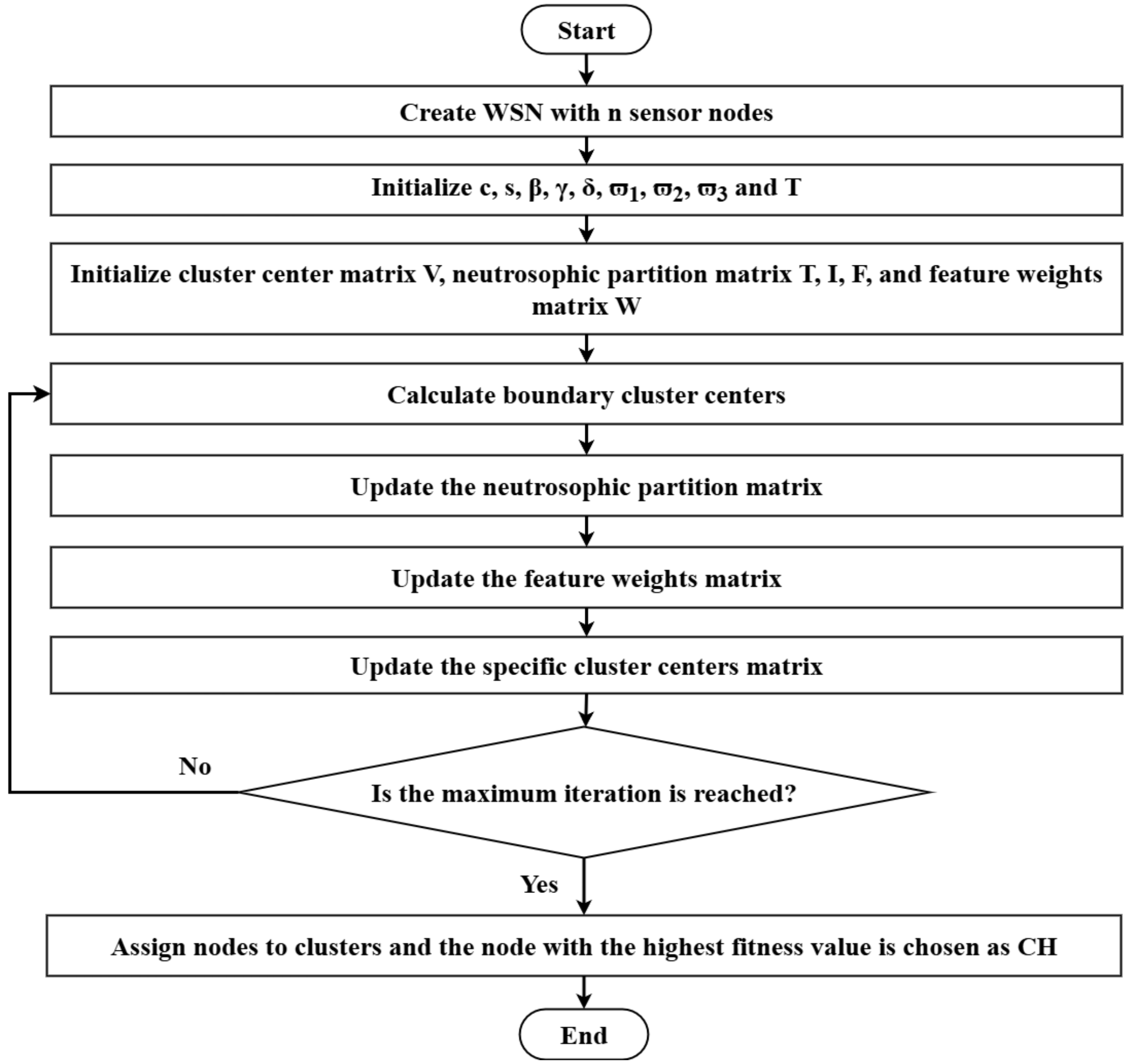

Figure 2 Flowchart of SWNCM Clustering Algorithm

Input: Sensor nodes $n = \{x_1, \dots, x_n\}$, number of clusters C, network factors s, β, γ, δ, ϖ_1, ϖ_2, and ϖ_3, and maximum iterations T

Output: Optimal clusters and CHs

1. Initialize V^0, T^0, I^0, F^0, and W^0;
2. while(maximum iteration is reached)
3. Calculate imprecise cluster centers utilizing equation (22);
4. Update the neutrosophic partition matrix utilizing equations (31)-(34);
5. Update feature weights utilizing equations (42)-(45);
6. Update specific cluster centers utilizing equation (36);
7. end while
8. Assign nodes to clusters;
9. For each cluster, choose the node that has the best fitness rating to serve as the CH;

Algorithm 1 SWNCM Clustering Algorithm

Table 2 Parameters for SWNCM Clustering

Parameters	Values
β	2
δ	$[10^{-2}, 10^{-1}, 1, 10^{1}, 10^{2}]$
$\varpi_1, \varpi_2, \varpi_3$	[0.01, 0.02, …, 0.98]
γ	$[2^{-5}, 2^{-4}, \dots, 2^{3}]$

RESEARCH ARTICLE

Thus, the SWNCM is performed for clustering the network in C clusters by optimizing the neutrosophic objective function defined in equation (20). Additionally, the node with the highest fitness function value, computed by equation (19), is elected for the best CHs in each cluster. Moreover, each CH switches dynamically across rounds, since SWNCM reclusters nodes and redetermines CHs in each round according to the updated fitness value. This demonstrates the adaptability of the presented SWNCM for fluctuations in network topology and node factors.

3.5. Greylag Goose Optimization Algorithm for Optimal Route Selection

At this point in the process, the CH has compiled all of CMs data and is ready to pass it on to the BS. To prolong network lifespan, it is essential to optimize multi-hop data transmission and choose the most efficient relay nodes for inter-cluster energy conservation while designing inter-cluster routing. When it comes to getting from the CHs to the BS, the GGO method is the way to go and the combinatorial optimization problem of selecting the best relay nodes is solved.

The GGO algorithm is selected for route optimization in this study due to its shown performance in similar optimization issues and unique features that could provide benefits over other optimization algorithms. The GGO starts by forming a group of people at random. For every optimization problem, everyone provides a path that might be considered optimal: the routes that carry data across clusters.

The initial inter-cluster path between all CHs and the BS is represented by n, and the GGO is comprised of people expressed as X_i, where $i = 1, \dots, n$. A fitness function, denoted by F_n, is chosen in order to assess the group members.

The GGO algorithm comprises two major processes: exploration and exploitation, which are balanced through iterative process to achieve efficient solutions.

3.5.1. Exploration

In the GGO algorithm, the exploration step is crucial; it involves searching the space for the likely best solution, that causes GGO to steer toward the global optima and avoid the local optimal.

Moving towards the ideal solutions: This tactic lowers the likelihood of becoming caught in local optima by using explores' geese, active seekers of optimal spots close to the existing ones. During iterations, the GGO technique uses the subsequent formulas $A = 2a \times r_1 - a$ to update vector A, $C = 2r_2$ to update C. From 2 to 0, the parameter a changes linearly:

$$X(t+1) = X^*(t) - A|C \times X^*(t) - X(t)| \times \tag{46}$$

In equation (46), the variables $X(t)$ and $X^*(t)$ stand for the positions of the ideal solution and each potential route at iteration t, respectively, while the integers r_1, r_2 range from 0 to 1. To improve exploration even further, this method incorporates the following equation and makes use of three randomly selected candidates (paddles): $X_{paddle\_1}, X_{paddle\_2}$ and X_{paddle_3}, this is further improved in equation (47):

$$X(t+1) = \omega_1 X_{paddle\_1} + z\omega_2 (X_{paddle\_2} - X_{paddle\_3}) + (1-z)\omega_3 (X - X_{paddle\_1}) \tag{47}$$

In equation (47), the weight values ω_1, ω_2 and ω_3 are within the interval [0,2]. The equation states that z declines gradually based on the t that represents $z = 1 - \left(\frac{t}{t_{max}}\right)^2$, where t as well as t_{max} are the iterations and upper limit, respectively. The next updating procedure includes reducing a and A for $r_3 \geq 0.5$:

$$X(t+1) = \omega_4 |X^*(t) - X(t)| e^{bl} \cos(2\pi l) + 2\omega_1 (r_4 + r_5) X^*(t) \tag{48}$$

In equation (48), b refers a constant, l refers to an arbitrary value in $[-1,1]$, ω_4 refers a weight value in [0,2], and r_4, r_5 denote arbitrary values in [0,1].

3.5.2. Exploitation

The algorithm's exploitative side is devoted to improving upon existing solutions. Once every iteration, GGO determines the entity demonstrating the highest level of fitness and awards appropriate recognition. There are two separate approaches to achieving exploitative goals, which are explained herein, and both contribute synergistically to improving the overall solution quality.

Approaching the optimal solution: In this process, three sentry solutions, $X_{Sentry\_1}, X_{Sentry\_2}$, as well as X_{Sentry_3} are used to move people $X_{NonSentry}$ towards the evaluated target position as given by the following equations (49) to (51):

$$X_1 = X_{Sentry\_1} - A_1 \times |C_1 \times X_{Sentry\_1} - X| \tag{49}$$

$$X_2 = X_{Sentry\_2} - A_2 \times |C_2 \times X_{Sentry\_2} - X| \tag{50}$$

$$X_3 = X_{Sentry\_3} - A_3 \times |C_3 \times X_{Sentry\_3} - X| \tag{51}$$

The variables A_1, A_2 and A_3 are defined as A in equations (49) to (51), while the variables C_1, C_2 and C_3 are defined as C. The population $X(t+1)$ and its simplified locations are given in equation (52):

$$X(t+1) = \frac{1}{3} \sum_{i=1}^{3} X_i \tag{52}$$

Searching the region around ideal solution: For improvement, each search areas close to the best leader, represented as X_{Flock_1} which is calculated by equation (53).

RESEARCH ARTICLE

$X(t+1) = X(t) + D(1+z)\omega\left(X - X_{Flock\_1}\right)$ (53)

$D, z,$ and ω are all variables that are used in the exploration technique in equation (53), where z is obtained using equation (54).

$$z = 1 - \left(\frac{t}{t_{max}}\right)^2 \quad (54)$$

In equation (54), t is the iteration number and t_{max} is the maximum iterations.

Choosing the best solution: This GGO incorporates a mutation approach and comprehensive assessment inside each group, suggesting an enhanced exploration capability of the algorithm. Updating the exploration (n_1) and exploitation (n_2) positions is a common component of each stage of the GGO method. Throughout the iterations, the parameter r_1, which is defined as $r_1 = c\left(1 - \frac{t}{t_{max}}\right)$, gets dynamically updated, with c being a constant. As each iteration comes to a close, GGO rearranges the candidates in the search space at random, changing their focus from exploration to exploitation. In the last stage, GGO offers the best option (the most effective transmission path). A pseudocode for entire process in this route optimization is presented in Algorithm 2.

Input: n sensor nodes, and required quantity of clusters C.

Output: Optimal transmission route between CHs and BS

1. n Set GGO $X_i, i = 1, \ldots, n$, population; maximum iterations t_{max}; fitness function F_n;
2. Assign initial values to the GGO parameters: $a, A, C, b, l, c, r_1, r_2, r_3, r_4, r_5, \omega, \omega_1, \omega_2, \omega_3, \omega_1, A_1, A_2, A_3, C_1, C_2, C_3$, and $t = 1$;
3. Determine the function of fitness F_n for every individual X_i (i.e., inter-cluster transmission route);
4. Fix P = best candidate location;
5. The exploration (n_1) as well as exploitation (n_2) solutions should be updated.
6. $\text{while}(t \le t_{max})$
7. $\text{for}(i1{:}\, i < n_1 + 1)$
8. $\text{if}(t\%2 == 0)$
9. $\text{if}(r_3 < 0.5)$
10. $\text{if}(|A| < 1)$
11. Use equation (46), revise the current candidate's position;
12. else
13. Choose three arbitrary candidates $X_{paddle\_1}, X_{paddle\_2}$, and X_{paddle_3};
14. Update z using equation (54);
15. Using equation (47), move the present potential solution to a different position;
16. end if
17. else
18. Use equation (49) to move the current candidate solution to a better spot;
19. end if
20. else
21. Revise the locations of the individuals using equation (53);
22. end if
23. end for
24. $\text{for}(i1{:}\, i < n_2 + 1)$
25. $\text{if}(t\%2 == 0)$
26. Determine X_1, X_2, and X_3 using equations (49) to (51);
27. Update individual locations as $x_i|_0^3$;
28. else
29. Update location of current candidate using equation (53);
30. end if
31. end for
32. For every X_i, find the fitness function;
33. Modify parameters;
34. Fixed $t = t + 1$;
35. Make adjustments to solutions outside of the search space;
36. if(best F_n is same as previous two results)
37. Improve n_1's exploration group's solutions while decreasing n_2's exploitation group's solutions;
38. end if
39. end while
40. Return optimal candidate P, i.e., optimal route between CH and BS;
41. End

Algorithm 2 GGO for Route Optimization

Thus, this study optimizes both CH and route selection process in WSNs with minimizing uncertainty for effective and reliable data transfer. Algorithm 3 presents an overall process in the presented INCRP for WSNs.

Input: n nodes

Output: Data transfer between source and BS

1. Begin
2. Construct the network with n nodes;
3. Determine the different network factors using equations (1) to (18);
4. Define the fitness function using equation (19);
5. Initialize C clusters and parameters for SWNCM;
6. Apply Algorithm 1 for clustering and CH selection;
7. Choose the CH for each cluster;
8. Initialize the GGO parameters;
9. Apply Algorithm 2 for optimal route selection;
10. Obtain the optimal path from CHs to BS;
11. CMs sense and transmit data to the respective CHs;
12. CHs aggregate data and forward it to the BS via the optimal path;
13. End

Algorithm 3 Proposed INCRP for WSNs

4. SIMULATION RESULTS

The proposed INCRP method for CH selection as well as data transfer in WSNs is evaluated in this section utilizing various performance metrics for its efficacy. The INCRP protocol is modelled in this study utilizing Network Simulator (NS3.26) on a Windows 10 64-bit OS, i7 processor, 8GB RAM, and 1TB HDD. Transmission latency, cluster stability, usage of energy, total residual energy, throughput, and network longevity are all metrics used for evaluation. The simulation configurations for both the suggested and conventional protocols are illustrated in Table 3.

Table 3 Simulation Parameters

Parameters	AMPRP-CIOA [17]	ISSA-C [18]	DPFCP [19] & HHOCFR [22]	ESR-HAC [24]	INCRP
	Range				
No. of sensor nodes	1000	50-1000	100	100-400	100-1000
E_{elec} nano Joule (nJ)/bit	-	-	50	50	50
Simulation area m^2	1000	1000	100	400	1000
No. of rounds	5000	1000	3000	1200	3000
Percentage of CHs	0.1	5	5	-	5
No. of search iterations	100	5	100	500	500
Data packet length (bits)	4000	4096	4000	2000	4096
Initial energy of each sensor node Joule(J)	0.25	0.5	1	0.5	0.5
ε_{mp} picoJoule (pJ)/bit/ m^2	-	-	0.0013	0.0013	0.0013
Distance threshold (d_0) in meters (m)	-	-	87.7	87	87.7
BS coordinates	-	-	(50, 50)	-	(500, 500)
ε_{fs} (pJ/bit/m^2)	-	-	10	10	10
Control packet size(bits)	-	-	200	-	200

RESEARCH ARTICLE

4.1. Performance Analysis of Proposed INCRP with Standard CRPs

The proposed INCRP is evaluated here in comparison to popular, well-established CRPs, including the, the Hybrid Energy-Efficient Distributed (HEED) clustering procedure, Threshold-sensitive Energy-Efficient Network (TEEN) protocol as well as Low Energy Adaptive Clustering Hierarchy (LEACH). From Figure 3 to Figure 9, the performance results were derived for 2500 rounds and 1000 nodes inside an area of 1000×1000 m². A 0.5 J early energy is assigned to each node, and the BS is located at (500, 500) coordinates. Data packets are thought to be 4096 bits long, whereas control packets are 200 bits long.

4.1.1. Throughput

This shows the total data received by the BS throughout the lifetime of the network. It is determined by equation (55).

$$\text{Throughput} = \frac{\text{Total received data (bits)}}{\text{Time (s)}} \quad (55)$$

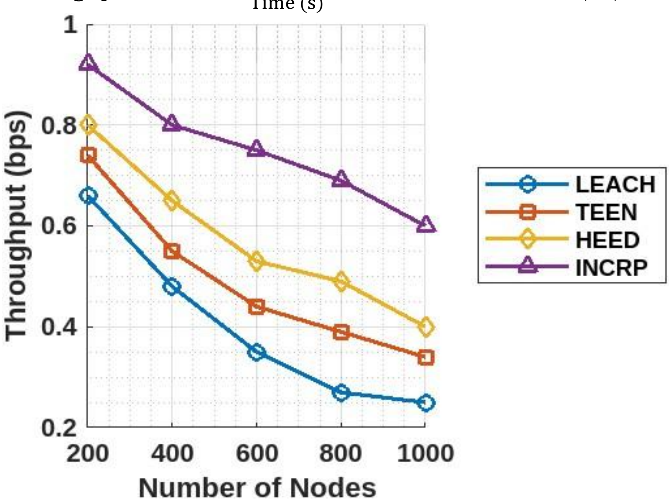

Figure 3 Evaluation of Throughput Vs Various Node Counts

The results of the throughput analysis for various node counts are displayed in Figure 3. It indicates that the INCRP for 200 nodes increases the throughput by 39.39%, 24.32%, and 15% compared to the LEACH, TEEN, and HEED protocols, respectively. As a result, it is revealed that the suggested INCRP improves throughput by selecting optimal CHs through the SWNCM and data transmission paths using the GGO algorithm. This SWNCM can choose the best CHs according to the different factors related to the node characteristics, energy, mobility, and traffic, which ensures the reliable data aggregation and transfer between the source and destination nodes.

4.1.2. Packet Delivery Ratio (PDR)

This metric measures how well packets are able to travel from their origin node to the BS on the network. It is calculated by equation (56).

$$PDR = \frac{\sum_1^n Total\ number\ of\ packets\ received}{\sum_1^n Total\ number\ of\ packets\ sent} \quad (56)$$

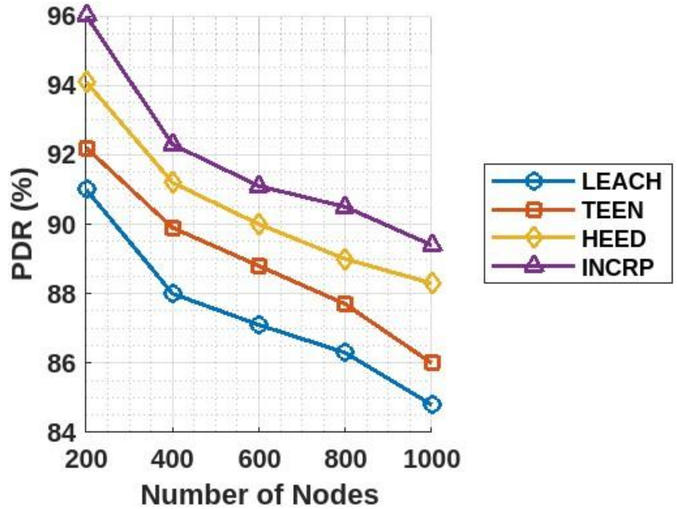

Figure 4 Comparison of PDR for Various Node Counts

The findings of the PDR assessment of performance for various nodes are displayed in Figure 4. It observes that the PDR of INCRP for 200 nodes is 5.49%, 4.12%, and 2.02% compared to the LEACH, TEEN, and HEED protocols, respectively. This improvement is achieved by adopting the SWNCM, which applies neutrosophic sets to perform clustering, resulting in reliable clustering and CH selection based on the different network factors, especially node density, traffic density, node centrality, and CH stability. These factors contribute to achieving a higher PDR in comparison to those protocols.

4.1.3. End-to-End Delay (E2D)

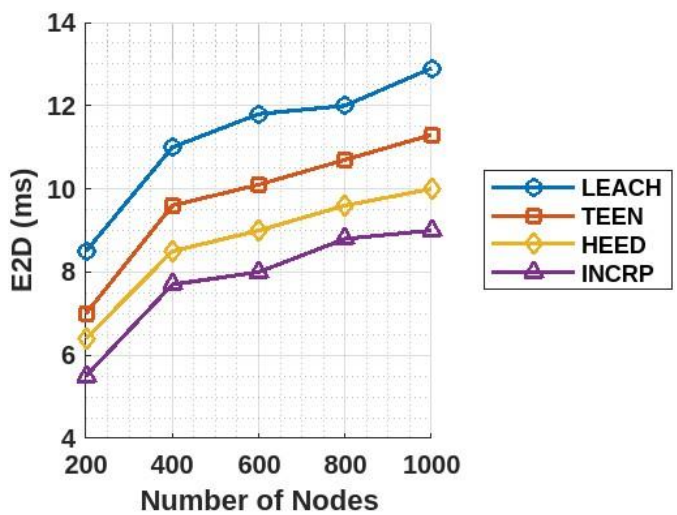

Figure 5 Comparison of E2D for Varying Node Counts

This has to do with how long it takes for data to go from sensor node to BS. It is determined by equation (57).

$$\text{E2D} = \text{processing delay} + \text{queuing delay} + \text{transmission delay} + \text{propagation delay} \quad (57)$$

Figure 5 demonstrates a comparison of E2D for different numbers of nodes. It indicates that the INCRP for 200 nodes reduces the E2D by 35.3%, 21.4%, and 14.1% compared to the LEACH, TEEN, and HEED protocols, respectively. This reduction is achieved due to the transfer of data packets from the source to the destination via the CHs chosen by the SWNCM, which utilizes delay as one of the input factors. Additionally, for inter-cluster data transfer, the GGO can select a less-congested, reliable multi-hop path based on delay as one of the fitness functions. Thus, the INCRP reduces the E2D significantly compared to the above-mentioned protocols.

4.1.4. Energy Consumption

The sum of all the energy consumed by sensor nodes during data reception, aggregation, and transmission is this value.

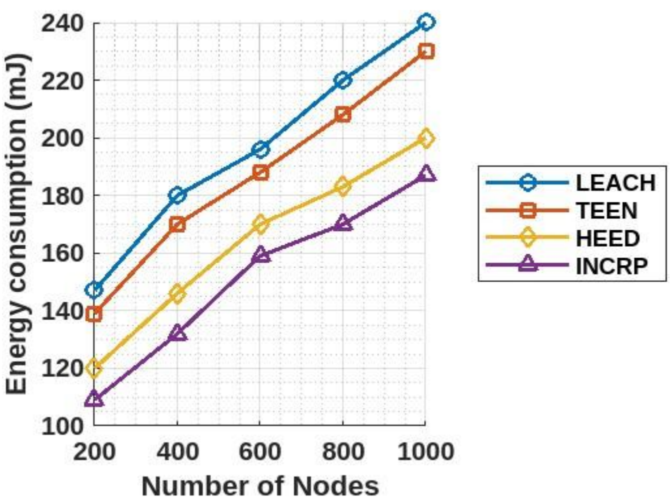

Figure 6 Comparison of Energy Consumption Varying Node Counts

Energy consumption (millijoules, or mJ) for different CRPs under varying node counts is analysed in Figure 6. Compared to the LEACH, TEEN, and HEED protocols, the energy usage of 200 nodes for the INCRP is lowered by 25.85%, 21.58%, and 9.17%, respectively. It reveals that the INCRP reduces the energy dissipation by considering energy as one of the network factors for choosing both optimal CHs by the SWNCM and a reliable path by the GGO. By considering energy, the INCRP can balance the energy dissipation to form the clusters and CH selection, as well as ensure the selection of energy-efficient routes for data transfer.

4.1.5. Alive Nodes

In this study, after a few rounds, the network efficiency is evaluated by looking at the number of alive nodes. The total number of sensor nodes with available energy is measured by this statistic. A longer lifetime for the network is indicative of a bigger number of live nodes.

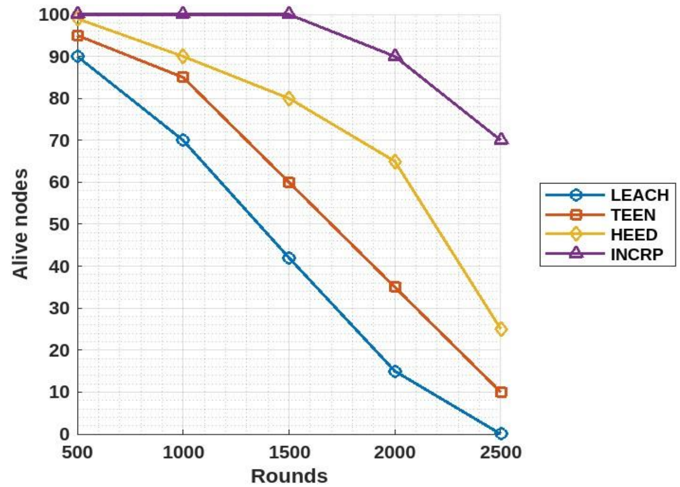

Figure 7 Comparison of Alive Nodes for Varying Rounds

Figure 7 presents the comparison of alive nodes for different rounds. For 2500 rounds, the alive node count for the suggested INCRP is 70, whereas the existing protocols, such as LEACH, TEEN, and HEED, have 0, 10, and 25 alive nodes. This implies that the proposed INCRP has more alive nodes after 2500 rounds, signifying a higher network lifespan. This improvement is due to the reduction in energy dissipation during clustering and CH selection, as well as the optimal path selection, which ensures energy saving across the network. This results in more alive nodes over varying rounds compared to the LEACH, TEEN, and HEED protocols.

4.1.6. Cluster Creation Time

It refers to the time taken to form clusters by partitioning the whole network.

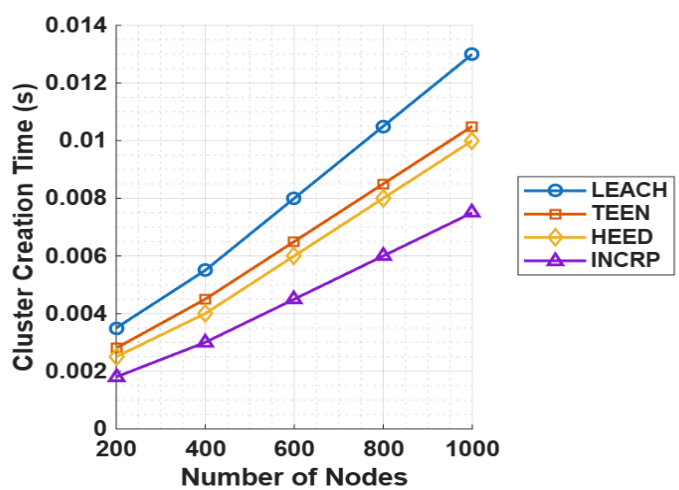

Figure 8 Comparison of Cluster Creation Time for Varying Nodes

RESEARCH ARTICLE

Figure 8 presents the comparison of cluster creation time with respect to varying node counts. It indicates that the cluster creation period of the suggested INCRP for 1000 nodes is reduced by 42.31%, 28.57%, and 25% compared to the LEACH, TEEN, and HEED, respectively. This reduction is due to the use of SWNCM's Lagrange optimization for effective partition and weight updates, which reduces the unnecessary computations in large-scale networks. Thus, the proposed INCRP lowers the cluster formation period.

4.1.7. CH Election Time

It refers to the average time taken to choose the best CHs for each cluster in the network.

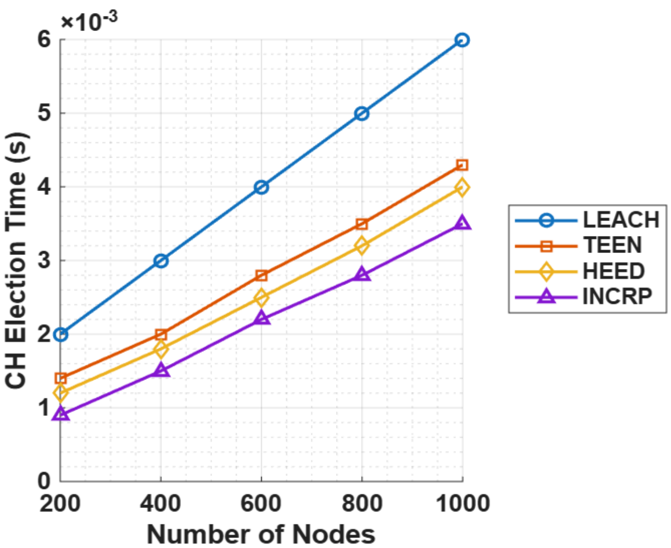

Figure 9 Evaluation of CH Election Timing for Different Nodes

Figure 9 presents the comparison of CH election time with respect to varying node counts. It observes that the CH election period of the suggested INCRP for 1000 nodes is reduced by 41.67%, 18.6%, and 12.5% compared to the LEACH, TEEN, and HEED, respectively. This reduction is achieved by adopting the neutrosophic objective function of SWNCM, which rapidly determines the best CHs based on various factors while precisely handling uncertainties. This leads to INCRP outperforming these protocols without additional computational overhead.

4.2. Ablation Studies

In comparison to traditional CRPs like AMPRP-CIOA [17], ISSA-C [18], DPFCP [19], HHOCFR [22] and ESR-HAC [24], and the outcomes of the performance assessment of the suggested INCRP are presented in this section. The outcomes of the comparison for the suggested INCRP are conducted across various scenarios, taking into account parameters like simulation area and the nodes count, among others.

- Scenario 1: Here, the suggested INCRP is contrasted with the previous AMPRP-CIOA [17] by employing a 1000×1000 m^2 area with 1000 nodes are deployed with starting energy of 0.25 nJ. There is a maximum number of 5,000 iterations for the simulation, and each data packet can only contain 4000 bits.
- Scenario 2: The INCRP is contrasted against the existing method ISSA-C [18] in this scenario, which employs 50-1000 sensor nodes with 0.5 J energy of each in a 1000×1000 m^2 area. With a data packet length of 4096 bits, the simulation can run for up to 1000 rounds.
- Scenario 3: The suggested INCRP is compared with the conventional DPFCP [19] and HHOCFR [22] in this scenario using 100 nodes in 100×100 m^2 area, with 1 J of initial energy. With a data packet length of 4000 bits, the simulation can run for up to 3000 rounds.
- Scenario 4: In this scenario, the existing ESAR-HAC [23] protocol is compared with the INCRP. In a 400×400 m^2 area, the simulation includes 100-400 sensor nodes 0.5 J energy. The results can be taken for 1200 rounds with 400 nodes and 800 rounds with 200 nodes. The data packet size is 2000 bits.

The suggested INCRP is compared to more traditional methods in a number of contexts in Table 4 to Table 9. This comparison analysis takes into account metrics like as energy consumption, throughput, First Node Dies (FND), alive nodes, PDR, lifetime of network, Half Node Dies (HND), as well as Last Node Dies (LND). In WSNs, the results show that the proposed INCRP outperforms in choosing the best CHs and routes for data transmission across different simulation rounds and numbers of sensor nodes.

Table 4 Comparative Analysis of Proposed INCRP vs AMPRP-CIOA

Network Area & Packet Size	Protocols	Metrics	No. of nodes				
			200	400	600	800	1000
1000×1000 m^2 & 4000 bits	AMPRP-CIOA [17]	Energy consumption (mJ)	0.9	1.7	1.3	1.2	1.4
		Network lifetime (×10^5)	2.84	3.00	2.94	2.98	3.10
		Throughput	0.95	0.86	0.94	0.88	0.92
	Proposed INCRP	Energy consumption (mJ)	0.84	0.98	1.15	1.07	1.25
		Network lifetime (×10^5)	3.15	3.32	3.48	3.60	3.73
		Throughput	0.97	0.90	0.96	0.93	0.95

Table 5 Comparative Analysis of Proposed INCRP vs ISSA-C

Network Area & Packet Size	Protocols	Metrics	No. of nodes				
			200	400	600	800	1000
1000×1000 m² & 4096 bits	ISSA-C[18]	PDR (%)	94.4	91.0	89.8	89.0	88.0
		Throughput (bps)	0.89	0.71	0.60	0.55	0.50
		Energy consumption (mJ)	120	150	170	190	210
		E2D (ms)	6.4	9.0	9.8	9.9	10.0
	Proposed INCRP	PDR (%)	96.0	92.3	91.1	90.5	89.4
		Throughput (bps)	0.92	0.80	0.75	0.69	0.60
		Energy consumption (mJ)	109	132	159	170	187
		E2D (ms)	5.5	7.7	8.0	8.8	9.0

Table 6 Evaluating the Suggested INCRP in Relation to DPFCP

Network Area & Packet Size	Protocols	Metrics	Rounds				
			500	1000	1500	2000	2500
100×100 m² & 4000 bits	DPFCP [19]	Energy consumption (J)	24	45	65	88	98
		Alive nodes	100	100	98	65	25
	Proposed INCRP	Energy consumption (J)	19	33	48	65	80
		Alive nodes	100	100	100	90	70

Table 7 Comparative Analysis of Proposed INCRP vs HHOCFR

Network Area & Packet Size	Protocols	Metrics	Rounds								
			200	400	600	800	1000	1200	1400	1600	1800
100×100 m² & 4000 bits	HHOCFR [22]	Energy consumption (J)	11	23	35	47	58	70	80	91	100
		Alive nodes	100	100	100	100	100	100	100	90	10
	Proposed INCRP	Energy consumption (J)	8	19	28	36	45	57	68	82	98
		Alive nodes	100	100	100	100	100	100	100	99	25

Table 8 Comparative Analysis of Proposed INCRP vs ESR-HAC

Network Area & Packet Size	Protocols	Metrics	No. of Nodes (Located in the Middle is the Sink Node)				No. of Nodes (Random Position in Perimeter)			
			100	200	300	400	100	200	300	400
400×400 m² & 2000 bits	ESR-HAC [24]	Packets received by sink ($\times 10^6$)	1.5	2.5	3.3	3.7	1.5	2.4	3.0	3.5
		PDR (%)	91.0	93.0	94.0	95.0	89.5	91.0	92.0	93.0
	Proposed INCRP	Packets received by sink ($\times 10^6$)	1.7	2.7	3.5	3.9	1.8	2.7	3.3	3.8
		PDR (%)	92.3	94.0	95.0	96.0	90.7	92.5	94.0	96.0

RESEARCH ARTICLE

Table 9 Comparative Analysis of Proposed INCRP vs ESR-HAC for Alive Nodes

Network Area & Packet Size	Protocols	Metrics (Alive Nodes)	No. of Nodes (The sink Node is Placed Centrally Within the Network)		No. of Nodes (Random Position in Perimeter)	
			200	400	200	400
400×400 m^2 & 2000 bits	ESR-HAC [24]	FND	637	850	561	832
		HND	709	962	665	912
		LND	732	1048	704	997
	Proposed INCRP	FND	703	980	679	954
		HND	756	1079	720	1048
		LND	781	1135	765	1113

4.2.1. Discussion

This section discusses the challenges of existing protocols and the merits of the suggested INCRP in WSNs. The AMPRP-CIOA [17] has low throughput when increasing the node count. The ISSA-C [18] has lower throughput and higher energy consumption. The existing DPFCP [19] has a higher energy consumption, and the PSO easily gets trapped in the local optima. The HHOCFR [22] has high energy consumption and requires other factors like node density and traffic density to further improve the network lifetime. The ESR-HAC [24] has a lower PDR, which may not be suitable for large and complex networks. These methods also rely entirely on variables like remaining power as well as distance to the BS. On the other hand, the proposed INCRP considers several factors, including residual energy, mean intra-cluster distance, mean BS distance, CH stability, node centrality, node degree, node density, traffic density, delay, and communication overhead for both CH and route selection in WSNs while handling uncertainty issues. This approach increases network performance in the case of energy consumption, alive nodes, E2D, PDR, and throughput compared to existing protocols.

Thus, the proposed INCRP can be well-suited for applications where network lifetime and scalability are vital. It can be applied in environmental monitoring, traffic control, industrial automation, battlefield monitoring, etc., since it sustains active nodes for a longer duration, ensuring effective data collection and transmission.

5. CONCLUSION

This study presented a novel INCRP to handle network uncertainty and extend the network lifespan in WSNs. First, the SWNCM clustering algorithm was used to cluster network nodes and pick the ideal CH based on factors such as mean distance between the BS and cluster member nodes, remaining energy, node density, node degree, latency, etc. Furthermore, the GGO method was utilized to decide the ideal transmission route from CHs to BS based on the same variables used in clustering. The findings presented that the INCRP has 89.4% PDR, a throughput of 0.6 bits per second (bps), consumed 187 mJ of energy, and has 99ms E2D for 1000 nodes in the network when compared to standard CRPs like LEACH, TEEN, and HEED. Future work will focus on detecting obstacles, node failures, or link failures when transmitting data across the network to further extend its longevity and energy efficiency.

REFERENCES

[1] K. Gulati, R. S. K. Boddu, D. Kapila, S. L. Bangare, N. Chandnani, and G. Saravanan, “A review paper on wireless sensor network techniques in Internet of Things (IoT),” Mater. Today Proc., vol. 51, pp. 161–165, Jan. 1, 2022.

[2] M. Majid, S. Habib, A. R. Javed, M. Rizwan, G. Srivastava, T. R. Gadekallu, and J. C. W. Lin, “Applications of wireless sensor networks and internet of things frameworks in the industry revolution 4.0: a systematic literature review,” Sensors, vol. 22, no. 6, p. 2087, Mar. 8, 2022.

[3] M. N. Hussain, M. A. Halim, M. Y. A. Khan, S. Ibrahim, and A. Haque, “A comprehensive review on techniques and challenges of energy harvesting from distributed renewable energy sources for wireless sensor networks,” Control Syst. Optim. Lett., vol. 2, no. 1, pp. 15–22, Jan. 29, 2024.

[4] A. I. Al-Sulaifanie, B. K. Al-Sulaifanie, and S. Biswas, “Recent trends in clustering algorithms for wireless sensor networks: a comprehensive review,” Comput. Commun., vol. 191, pp. 395–424, Jul. 1, 2022.

[5] S. Kumar and A. Shankar, “Cluster-based energy-efficient routing protocol in next generation sensor networks,” Int. J. Grid Util. Comput., vol. 15, no. 2, pp. 181–197, Apr. 8, 2024.

[6] P. Karpurasundharapondian and M. Selvi, “A comprehensive survey on optimization techniques for efficient cluster based routing in WSN,” Peer-to-Peer Netw. Appl., vol. 17, no. 5, pp. 3080-3093, Sep. 2024.

[7] M. K. Reddy, A. S. Veluswamy, S. Selvanayaki, C. Harini, P. Kumar, and S. Shameem, S. “Modeling of metaheuristic-based dual cluster head selection with routing protocol for energy-efficient wireless sensor networks,” Soft Comput., vol. 29, no. 6, pp. 2999-3020, Mar. 2025.

[8] C. Del-Valle-Soto, A. Rodríguez, and C. R. Ascencio-Piña, “A survey of energy-efficient clustering routing protocols for wireless sensor networks based on metaheuristic approaches,” Artif. Intell. Rev., vol. 56, no. 9, pp. 9699–9770, Sep. 2023.

[9] S. El Khediri, A. Selmi, R. U. Khan, T. Moulahi, and P. Lorenz, “Energy efficient cluster routing protocol for wireless sensor networks

RESEARCH ARTICLE

using hybrid metaheuristic approaches," Ad Hoc Netw., vol. 158, p. 103473, May 1, 2024.
[10] S. El Khediri, W. Fakhet, T. Moulahi, R. Khan, A. Thaljaoui, and A. Kachouri, "Improved node localization using K-means clustering for wireless sensor networks," Comput. Sci. Rev., vol. 37, p. 100284, Aug. 1, 2020.
[11] M. Adnan, L. Yang, T. Ahmad, and Y. Tao, "An unequally clustered multi-hop routing protocol based on fuzzy logic for wireless sensor networks," IEEE Access, vol. 9, pp. 38531–38545, Mar. 2, 2021.
[12] H. Hu, X. Fan, and C. Wang, "Energy efficient clustering and routing protocol based on quantum particle swarm optimization and fuzzy logic for wireless sensor networks," Sci. Rep., vol. 14, no. 1, p. 18595, Aug. 10, 2024.
[13] L. Sahoo, S. S. Sen, K. Tiwary, S. Moslem, and T. Senapati, "Improvement of wireless sensor network lifetime via intelligent clustering under uncertainty," IEEE Access, vol. 12, pp. 25018–25033, Feb. 13, 2024.
[14] 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. 2, 2023.
[15] A. Jalili, M. Gheisari, J. A. Alzubi, C. Fernández-Campusano, F. Kamalov, and S. Moussa, "A novel model for efficient cluster head selection in mobile WSNs using residual energy and neural networks," Meas.: Sens., vol. 33, p. 101144, Jun. 1, 2024.
[16] R. A. Pravin, K. Murugan, C. Thiripurasundari, P. R. Christodoss, R. Puviarasi, and S. I. A. Lathif, "Stochastic cluster head selection model for energy balancing in IoT enabled heterogeneous WSN," Meas.: Sens., vol. 35, p. 101282, Oct. 1, 2024.
[17] P. Divya and B. Sudhakar, "Route optimization and optimal cluster head selection for cluster-oriented wireless sensor network utilizing circle-inspired optimization algorithm," Int. J. Comput. Intell. Syst., vol. 17, no. 1, p. 302, Dec. 18, 2024.
[18] G. H. Alshammri, "Enhancing wireless sensor network lifespan and efficiency through improved cluster head selection using improved squirrel search algorithm," Artif. Intell. Rev., vol. 58, no. 3, p. 79, Jan. 6, 2025.
[19] C. Wang, "A distributed particle-swarm-optimization-based fuzzy clustering protocol for wireless sensor networks," Sensors, vol. 23, no. 15, p. 6699, Jul. 26, 2023.
[20] L. Yuebo, Y. Haitao, L. Hongyan, and L. Qingxue, "Fuzzy clustering and routing protocol with rules tuned by improved particle swarm optimization for wireless sensor networks," IEEE Access, vol. 11, pp. 128784–128800, Nov. 14, 2023.
[21] M. Manoharan, B. Subramani, and P. Ramu, "An optimal energy efficient routing in WSN using adaptive entropy bald eagle search optimization and density based adaptive soft clustering," Sustain. Comput. Inform. Syst., vol. 43, p. 101003, Sep. 1, 2024.
[22] D. Jing, "Harris harks optimization based clustering with fuzzy routing for lifetime enhancing in wireless sensor networks," IEEE Access, vol. 12, pp. 12149–12163, Jan. 15, 2024.
[23] R. Chen, "Optimizing wireless sensor network topology with node load consideration," Virtual Real. Intell. Hardw., vol. 7, no. 1, pp. 47–61, Feb. 1, 2025.
[24] X. Chai, Y. Wu, and L. Feng, "Energy-efficient scalable routing algorithm based on hierarchical agglomerative clustering for wireless sensor networks," Alexandria Eng. J., vol. 120, pp. 95–105, May 1, 2025.

Authors

Mr. C.A. Kandasamy has completed his schooling at Government Boys Higher Secondary School, Tiruchengode and graduated in 2008 with a degree in B.C.A. and Master of Computer Applications in 2011 at K. S. Rangasamy College of Technology, Tiruchengode. He obtained his Part Time M.Phil. degree in Computer Science from Bharathiar University, Coimbatore in 2014. He was Currently Pursing with part time (Ph.D.) in the area of Networking from K.P.R College of Arts Science and Research affiliated to Bharathiar University, Coimbatore. He was joined as an Assistant Professor in MCA Department at K.S.R College of Engineering, Tiruchengode from 2011 to till date. He was 14 years of Teaching Experience. He published around 10 research papers in National, International Journals and Conferences. His areas of interests are Software Engineering, Cloud Computing, Big Data and Analytics, Networking.

Dr. D. MAHESHWARI is currently working as an Associate Professor and Head, Department of Computer Technology, KPR College of Arts Science and Research, Coimbatore. She received her Ph.D. degree in Computer Science from Bharathiar University, Coimbatore. He received her master degree in M.Sc. Computer Science and from bachelor degree in B.Sc. Computer Science from Kovai Kalaimagal College of Arts and Science, Coimbatore. Having 20+ years of experience in Teaching & Research. Her area of research includes Wireless Network Security and Edge Computing. She is as Reviewer of 6 National/International Journals and Conferences. As inspired writer, she has authored 19 Research Papers, 2 Books and 35 Technical Presentation. She received honors and awards from various institutions.

How to cite this article:

C.A. Kandasamy, D. Maheshwari, "An Intelligent Neutrosophic Clustering based Optimized Routing Protocol for Wireless Sensor Networks", International Journal of Computer Networks and Applications (IJCNA), 12(5), PP: 809-827, 2025, DOI: 10.22247/ijcna/2025/48.