Received: May 3, 2024. Revised: June 8, 2024.

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Distributed Cluster Head Selection Using Progressive Index-Modulation for Energy-Efficient Inter and Intra-Cluster Communication in Multi-Hop WSNs

Monika Parameswaran[1]* **Vijayalakshmi Shanmugam[2]**

[1]*Department of Computer Science, Sri Ramakrishna College of Arts and Science for Women, Coimbatore - 641004, Tamil Nadu, India*
[2]*Department of Data Analytics, Sri Ramakrishna College of Arts and Science for Women, Coimbatore - 641004, Tamil Nadu, India*
*Corresponding author's Email: monika19594phd@gmail.com

Abstract: Energy minimization is crucial in designing Wireless Sensor Networks (WSNs), especially for multi-hop transmission. To address this challenge, an Intra-Cluster and Inter-Cluster (ICIC) energy reduction mechanism was developed, which includes dynamic duty cycle allocation for intra-cluster transmission and selection of the lowest Energy Consumption (EC) path for inter-cluster transmission. Conversely, randomly selecting Cluster Heads (CHs) can lead to inefficient distribution of tasks among nodes, impacting network lifetime. Additionally, small errors in the data received can lead to the Sink Node (SN) making an unsuccessful global decision in different WSN applications. Therefore, this manuscript introduces a new Progressive Index-Modulation (PIM)-based CH selection algorithm in ICIC (PIM-ICIC) for multi-hop WSNs. In this algorithm, a novel broadcasting scheme called index-shift is adopted to distribute the CH role more evenly among sensor nodes. First, Cluster Member (CM) nodes are divided into sets where nodes within the same set collaborate in selecting CHs to balance the CH loads. The CH is chosen according to the local cluster data in each cycle. If a node reaches a threshold limit for being chosen as a CH, an alternative node from a similar set can assume the CH role to transmit data to the SN. Thus, the PIM-ICIC algorithm can prolong network lifetime and reduce packet loss effectively. Finally, extensive simulations demonstrate that the PIM-ICIC achieves a Packet Delivery Ratio (PDR) of 5.8×10^5 packets, EC of 14J, and 20 dead nodes over 1000 rounds. With 400 nodes in the network, it achieves a PDR of 6.9×10^5 packets, 78% EC, and 86 dead nodes. At a 40dB Signal-to-Noise Ratio (SNR), PIM-ICIC achieves a Data Error Rate (DER) of 0.0002, surpassing ICIC, Hybrid Particle Swarm Optimization with Improved Low Energy Adaptive Clustering Hierarchy (HPSO-ILEACH), and Sine Cosine Algorithm with Levy mutation (SCA-Levy) in multi-hop WSNs.

Keywords: Multi-hop WSN, Clustering, Energy efficiency, Inter-cluster transmission, Intra-cluster transmission, Duty cycle, Index modulation.

1. Introduction

WSNs are widely utilized in industries, home automation, atmosphere monitoring, etc., due to their compactness, affordability, and ease of use [1]. In environments where energy depletion is a concern, such as mining, extending the network's lifespan is crucial as batteries are not easily replaceable or rechargeable. Additionally, the size of the deployed area significantly impacts WSN design. In a limited area, information is forwarded directly from the node to the SN or Base Station (BS). In a wide area, data is relayed to the SN via multiple intermediary nodes [2]. Constructing a broad area network involves three main issues: subdividing the region, optimizing routing, and managing the energy of intermediary nodes [3]. The entire zone is divided into smaller sub-zones called clusters, with one node designated as the CH responsible for accepting information from CMs and transmitting it to the SN or subsequent node [4].

Various clustering schemes including the Low Energy Adaptive Clustering Hierarchy (LEACH) scheme, etc., are employed in WSNs to prolong network lifetime and optimize EC [5, 6].

Routing strategies in WSNs include flat routing, establishing a direct route to the SN [7], and hierarchical routing, where data is transmitted from sensor nodes to CH nodes and then forwarded to the next CH node [8]. Additionally, relay nodes play a critical role in multi-hop transmission by ensuring that intermediate nodes have sufficient energy for relaying packets [9]. Thus, optimal route selection is essential for efficient WSN design to prevent premature energy depletion and the formation of dead node regions near the SN [10]. A sensor node in a WSN comprises a sensor, radio transmitter, receiver, processing unit, and battery. These components utilize energy in various ways, including data sensing, processing, and signal transmission. Sensing energy is used to detect physical quantities like temperature, gas, and humidity, requiring less energy compared to other components. The processing unit processes sensor data, converts it into packets for transmission, and handles data received from other nodes. The radio transmitter and receiver consume the most energy in the node, with EC increasing with the distance of transmission in the radio model [11].

From this perspective, Shah et al. [12] address the issue of EC in multi-hop WSNs. They consider routing, relay nodes, and wormhole issues during WSN design. They proposed the ICIC energy minimization strategy in multi-hop WSNs. In intra-cluster transmission, a node with sufficient energy is selected as the CH for data gathering and transmission. Other nodes send information to the CH. EC is decreased by adjusting the node’s transfer period. Nodes closer to the CH node are given more transmission time, resulting in more packets delivered with equal remaining energy than the remote nodes. The active period if a node forwards data and the passive time when it does not are balanced using duty cycling. Various Medium Access Control (MAC) protocols manage the on/off transmission time to improve energy efficiency [13]. The Time Division Multiple Access (TDMA) scheme is a suitable selection in WSNs. In inter-cluster communication, CHs forward information to the next CH until all packets reach the SN. The route requiring the lowest energy for packet transfer is elected for inter-cluster communication.

1.1 Problem statement

In the ICIC protocol, the CH selection is done randomly, which poses a challenge as the CH functionality needs to be distributed resourcefully among nodes to lengthen the network lifespan. To address this issue, the CH selection process incorporates criteria such as nodes' energy levels, connectivity between nodes, and distance from the SN. Hybrid schemes combine various criteria using optimization methods to choose the best CH. The selected CH gathers information from nodes and forwards it to the SN. However, small errors in the data received can lead to the SN making an unsuccessful global decision in different WSN applications. To improve data accuracy, Index Modulation (IM) schemes are used in WSNs [14], where part of the CH data is transmitted through the index of the CH, and the rest is encoded in a particular character and sent to the SN. The SN decodes the received data using the Maximum Likelihood Estimator (MLE). Though this scheme reduces the data error, it has a high difficulty for the SN owing to the brute-force exploration needed by the MLE.

1.2 Major contributions of the paper

This manuscript introduces a new PIM-based CH selection algorithm for multi-hop WSNs called the PIM-ICIC algorithm. The goal is to enhance the CH selection process used in IM broadcasting schemes. A novel broadcasting method, called index-shift, is introduced to distribute the CH role more evenly among sensor nodes. First, CMs are divided into sets where nodes within the same set collaborate in selecting CHs to balance the CH loads. The CH is chosen according to the local cluster data in each cycle. If the node reaches a threshold limit for being chosen as a CH, an alternative node from a similar set can assume the CH role to transmit data to the SN. Thus, the PIM-ICIC algorithm can prolong network lifetime and reduce packet loss effectively. Finally, extensive simulations demonstrate that the PIM-ICIC yields superior network performance compared to existing algorithms.

The remaining sections include: Section 2 covers the literature survey. Section 3 describes the PIM-ICIC algorithm and Section 4 evaluates its performance against earlier ones. Section 5 précises the study.

2. Literature survey

This section discusses recent clustering algorithms in WSNs. A new energy-aware CH selection algorithm [15] was developed to select CHs in WSNs according to aspects like remaining power, node position, and centrality. However, more dead nodes were occurred due to high EC when the distance between CH and SN is high. This leads to high DER and lower PDR. A correlation model [16] was developed to choose nodes from each correlated region for clustering, while the rest entered sleep mode. They also introduced a Fuzzy-based

Received: May 3, 2024. Revised: June 8, 2024.

distributed cluster technique to address uncertainty issues. However, achieving a low DER and high PDR was challenging due to the scalability issues in data-gathering process with larger network sizes.

A clustering protocol [17] was implemented for WSN using Fuzzy Logic (FL) and multi-hop transmission. Input variables for the CH selection include node distance to the BS, concentration, and residual energy. However, the number of dead nodes was high due to congestion between clusters and the rapid depletion of nodes near the sink. A distributed multi-level clustering algorithm using Sugeno-based FL controller and an enhanced Squirrel Search Algorithm (SSA) [18] was presented to lessen EC in WSNs. However, the PDR and EC was influenced by the CH failures since the reliability of the clustering was not taken into account.

An Energy-Efficient CH selection utilizing the Improved Grey Wolf Optimization (EECHIGWO) scheme [19] was developed, which optimizes CH choice based on parameters like sink distance, remaining power, CH balancing variable, and mean intra-cluster distance. However, the PDR and number of dead nodes were impacted by the limited network size. The Cluster-Centered CH selection Algorithm (C3HA) [20] was developed using k-means to reduce EC in WSNs and extend their lifespan. However, the PDR and DER were not satisfactory. The HPSO-ILEACH [21] algorithm was developed for WSNs. The HPSO was first utilized to choose the optimal CHs based on factors like distance between cluster members and remaining node energy. Subsequently, ILEACH was used to minimize EC by adjusting the CHs in the clustering process. However, the DER and SNR was high due to the random distribution of CHs.

The SCA-Levy based WSN clustering routing scheme [22] was developed that integrates the SCA and Levy mutation. The SCA was utilized to elect CHs, while levy mutation was employed to enhance global exploitation ability. A selection method based on remaining energy and intra-cluster distance was introduced. Also, optimal relay nodes were selected based on their position and the residual energy of the BS to improve longevity and reduce the need for long-range communication. However, the DER was high since the random distribution of CHs affect the data gathering process.

2.1 Research gap

The literature highlights that the clustering and CH choice vary in criteria. However, errors in data received by the SN from CHs can hinder global decision-making in WSN applications. To address this, a novel PIM scheme is proposed in this study to balance the CH role among nodes and minimize data errors at the SN.

3. Proposed methodology

This section provides a detailed explanation of the PIM-ICIC algorithm in multi-hop WSNs. It includes descriptions of network and energy models, as well as a brief overview of CH selection using the PIM scheme. Table 1 lists the notations used in this study.

3.1 Network model

Consider a WSN with N nodes organized into C clusters according to vicinity, each containing N_c nodes, as illustrated in Fig. 1. Each node monitors target activity through a 2-state Bernoulli process with possibilities P_1 and P_0 for Target Presence (TP) and Target Absence (TA), respectively. Each node makes a local decision $e_n = \{0,1\}$ indicating TA or TP, where $0 \leq n \leq N_c - 1$. The accurateness of these local decisions is characterized by 2 discovery possibilities: λ_{ii} and λ_{ij}, where $i, j \in \{0,1\}$. λ_{ii} represents the possibility of appropriately detecting the target position, and λ_{ij} denotes the possibility of erroneously deciding the target position. Both λ_{ii} and λ_{ij} are contingent on the sensing excellence of the CMs and are believed to be equal across each node.

Local decisions from nodes are sent to the CH, which forwards them to the SN. The SN processes these decisions to make a global decision about TP according to the majority rule. When the amount of positive decisions meets a threshold K, the target is considered present. The SN coordinates

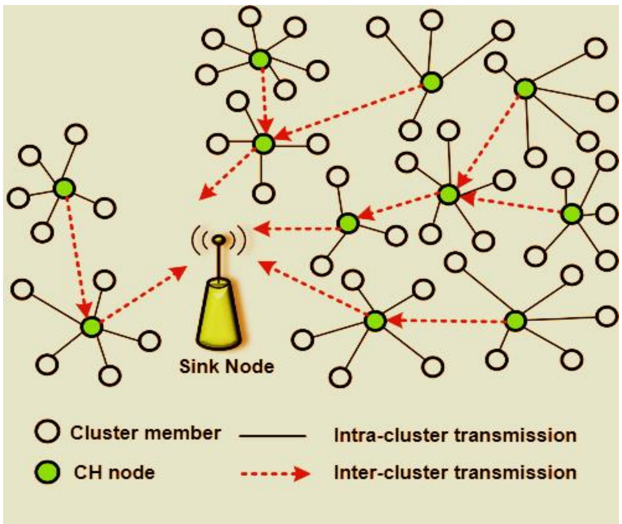

Figure. 1 Structure of cluster-based multi-hop WSN

Table 1. Lists of notations

Notations	Description
N	No. of nodes
C	No. of clusters
N_c	No. of nodes in each cluster
P_1 and P_0	Possibilities for TP and TA, respectively
e_n	Local decision
λ_{ii}	Possibility of appropriately detecting the target position
λ_{ij}	Possibility of erroneously deciding the target position
K	Threshold
y	Received signal at the SN
H	Channel matrix with Rayleigh fading model
N_r	Amount of receiving antennas at the SN
w	Received noise vector
σ^2	Standard variance
x	Communication vector
E_{Rec_CM}	Energy used by CM for receiving data
E_{sense}	Sensing energy needed by CM node
P_{CH}	No. of packets transferred from the nearest CH
b	No. of bits
P_{CM}	Amount of packets accepted by a CH from CMs
E_{Rec_CH}	Energy used by the CH for receiving P_{CH}
$E_{totalRec\_CH}$	Total EC in receiving data packets
d	Distance
$E_{totaltrans\_CM}$	Energy needed to transferring packets to CH from CM at d
E_{elec}	Energy used in the transmitter to process a bit
d_0	Threshold distance
ϵ_{fs} and ϵ_{mp}	Free-space and multi-path amplification coefficients, respectively
$E_{totaltrans\_CH}$	Energy used for transferring packets from CH to the successive CH
$\hat{\alpha}$	Node index
S	No. of CM node sets
N_S	No. of nodes in each set S
θ_i	CH counter
θ_{max}	Maximum value of CH counter
$\hat{\beta}$	Modulated data
r	Index-shift
$\hat{x}$	Cluster information
$e_0, \dots, e_{\log_2 N_c - 1}$	Cluster information bits
l	Length of cluster information bits
α	CH index
k	Amount of 1s in the cluster information bits
$\mathcal{P}_{PIM}(i)$	Possibility of allocating a node with index i as a CH
$\bar{\mathcal{P}}_{index-shift}(i)$	Mean possibility of electing i as the CH for $\theta_{max} < \infty$
$\mathcal{P}_F$ and $\mathcal{P}_D$	Global false alarm possibility and the global discovery possibility, respectively
δ	Amount of 1s in the transferred cluster information
δ'	Amount of 1s in the accepted cluster information
$\chi^{(i)}$ and $\chi^{(\delta'_m)}$	Sets containing potential cluster information with sums i and δ'_m, respectively
$Pr.\{x^{(p)} \to x^{(q)}\}$	Probability of receiving vector $x^{(q)}$ given that the transmitted vector is $x^{(p)}$
γ	Gamma random factor

communications from clusters using a TDMA technique with dedicated time slots. The received signal at the SN is represented by Eq. (1).

$$y = Hx + w \tag{1}$$

In Eq. (1), H is the channel matrix $(N_r \times N_c)$ with Rayleigh fading model, N_r represents the amount of receiving antennas at the SN, w denotes the received noise vector $(N_r \times 1)$ with complex Gaussian random factors of 0 mean and σ^2 variance, and x represents the communication vector $(N_c \times 1)$ with transmitted signals from each node. Each element of x is mostly zeros except for the CH element, which equals the modulated symbol. The channel between the CMs and the SN is designed as a dispersed Multi-Input Multi-Output (MIMO) channel.

3.2 Energy model

A node consumes energy when transferring or receiving data. Data can be received from CM nodes or nearby CHs and transferred from CM nodes to CHs or from CHs to the next CHs. Let P_{CM} represent the sum amount of packets accepted by a CH from CMs, all containing b bits.

The energy used by each CM node for receiving data $\left(E_{Rec\_CM}\right)$ is calculated by Eq. (2).

$$E_{Rec\_CM} = E_{sense} \times P_{CM} \times b \tag{2}$$

In Eq. (2), E_{sense} represents the sensing energy required by CM nodes. Let P_{CH} represent the number of packets transferred from the nearest CH. The energy used by the CH for receiving P_{CH} $\left(E_{Rec\_CH}\right)$ is calculated by Eq. (3).

$$E_{Rec\_CH} = E_{sense} \times P_{CH} \times b \tag{3}$$

Therefore, the total EC in receiving data packets $\left(E_{totalRec\_CH}\right)$ is calculated by Eq. (4).

$$E_{totalRec_{CH}} = E_{Rec_{CM}} + E_{Rec_{CH}} = E_{sense} \times b(P_{CM} + P_{CH}) \tag{4}$$

The energy needed to transferring packets to CH from CM nodes at distance d [23], $\left(E_{totaltrans\_CM}\right)$ is calculated by Eqns. (5) & (6).

$$E_{totaltrans\_CM} = \begin{cases} b \times P_{CM}\left(E_{elec} + \epsilon_{fs} \times d^2\right), & d < d_0 \\ b \times P_{CM}\left(E_{elec} + \epsilon_{mp} \times d^4\right), & d \geq d_0 \end{cases} \tag{5}$$

$$\text{Where } d_0 = \sqrt{\epsilon_{fs}/\epsilon_{mp}} \tag{6}$$

In Eq. (5), E_{elec} denotes the energy used in the transmitter to process a bit, d_0 denotes the threshold distance, and $\epsilon_{fs}, \epsilon_{mp}$ represent free-space and multi-path amplification coefficients, respectively. Likewise, the energy used for transferring packets from CH to the successive CH $\left(E_{totaltrans\_CH}\right)$ is calculated by Eq. (7).

$$E_{totaltrans\_CH} = \begin{cases} b \times P_{CH}\left(E_{elec} + \epsilon_{fs} \times d^2\right), & d < d_0 \\ b \times P_{CH}\left(E_{elec} + \epsilon_{mp} \times d^4\right), & d \geq d_0 \end{cases} \tag{7}$$

Therefore, the total energy used for transmitting packets $\left(E_{totaltrans\_CH}\right)$ is calculated by Eq. (8).

$$E_{totaltrans_{CH}} = E_{totaltrans\_CM} + E_{totaltrans\_CH} \tag{8}$$

3.3 Cluster head selection

This study uses an index-shift scheme to select CHs in multi-hop WSNs, considering local cluster information and CH choice history. Nodes are grouped to evenly distribute CH load, with CH selection according to the local information. If a node's CH counter touches a threshold, alternative node from a similar set takes over the CH part for efficient data transmission to the SN.

The PIM-ICIC algorithm is extended to transmit the cluster information using the CH's indices, which are calculated by the total entire cluster information. Nonetheless, there are $N_c + 1$ probable indices $(0, \dots, N_c)$, which is greater than the accessible indices N_c. To address this issue, nodes will remain silent (no signal transmission) if the information sum is 0, as there is no need to report when no targets are present.

The CH index is chosen according to the node indices $\left(0, \dots, (N_c - 1)\right)$ and the information totality values $(1, \dots, N_c)$ as Eq. (9).

$$\hat{\alpha} = \left(\sum_{n=0}^{N_c-1} e_n\right) - 1 \tag{9}$$

The chosen CH node with index $\hat{\alpha}$ will inform the SN regarding the local cluster information. The CH index fluctuates arbitrarily according to the overall cluster information, leading to an unbalanced distribution among cluster nodes. To balance the CH part, CMs are organized into S sets $(v_0, \ldots, v_{S-1})$, where each containing $N_S = \frac{N_c}{S}$ nodes.

Nodes within the same set coordinate to evenly distribute the CH load. A node j in set v_i is elected as the CH based on prior CH counts, given a specific $\hat{\alpha}$ in v_i. To choose the CH within a set, a CH counter θ_i is assigned to each node $i = 0, \ldots, N_c - 1$.

Primarily, θ_i is assigned to 0 and disseminated among nodes in a similar set. When the CH counter of node i reaches a fixed highest value θ_{max}, alternative node with index r, different from i, becomes the CH as long as its CH counter is less than θ_{max}. This process continues until each node in the set have attained θ_{max}. Then, θ is reset for these nodes and the process repeats. All nodes in the set are aware of the current CH selection and modify their CH counter replicas θ_i.

The chosen CH can convert the index-shift r into a particular sign $\hat{\beta}$ and send it to the SN. The sign is converted using N_S-signal modulation order, where $r = 0,1, \ldots, N_S - 1$. $\hat{\beta}$ is calculated by Eq. (10).

$$\hat{\beta} = \mathrm{mod}_{N_S}\{r\} \tag{10}$$

Each cluster transmits a vector x with zeros in all elements except for the active CH index containing the modulated sign. The SN has N_S potential r for all clusters, resulting in $N_c \times N_S$ potential solutions for all received signs. The SN uses a MLE, as defined in Eq. (11), to search for the cluster information by considering every potential result.

$$\hat{x} = \underset{i=1:2^{N_c}}{\mathrm{argmin}} \left\| y - Hx^{(i)} \right\|^2 \tag{11}$$

3.4 Distribution of CH role

This subdivision analyzes the distribution of CH roles in different reporting patterns. The probabilities λ_{ii} and λ_{ij} are influenced by false-alarm λ_{01} and miss-detection λ_{10} probabilities. It is important to note that these probabilities are assumed to be the same for all nodes. In the PIM-ICIC algorithm, the cluster information bits $e_0, \ldots, e_{\log_2 N_c - 1}$ correspond to selecting the CH index α as shown in Eq. (12).

$$\alpha = dec\{e_0, e_1, \ldots, e_{b-1}\} \tag{12}$$

Consider that the length of these bits is represented by l, where $l = \log_2 N_c$. So, the possibility of allocating a node with index i as a CH for a cluster information is determined by Eq. (13).

$$\mathcal{P}_{PIM}(i) = \lambda_{00}^{(l-k)} \lambda_{01}^{k} \mathcal{P}_0 + \lambda_{11}^{k} \lambda_{10}^{(l-k)} \mathcal{P}_1, i = 0,1, \ldots, N_c - 1 \tag{13}$$

In Eq. (13), k represents the amount of 1s in the bits $e_0, \ldots, e_{l-1}$. However, for the PIM-ICIC, the total cluster information bits $e_0, \ldots, e_{N_c-2}$ are the CH index $\tilde{\alpha}$ is provided in Eq. (14).

$$\tilde{\alpha} = \sum_{n=0}^{N_c-2} e_n \tag{14}$$

The length of these bits is $N_c - 1$. Therefore, the possibility of allocating i as a CH for a cluster information is determined by Eq. (15).

$$\mathcal{P}_{PIM}(i) = \left(\lambda_{00}^{(N_c-i-1)} \lambda_{01}^{i} \mathcal{P}_0 + \lambda_{11}^{i} \lambda_{10}^{(N_c-i-1)} \mathcal{P}_1 \right) \binom{N_c - 1}{i} \tag{15}$$

In the index-shift scheme, the possibility of electing i as the CH follows $\mathcal{P}_{PIM}(i)$ as given by Eq. (15). The sum of the entire cluster information $e_0, \ldots, e_{N_c-1}$ is deliberated for the CH index choice excluding cases where the cluster information total is 0. Thus, the possibility of allocating i as the CH is provided by Eq. (16).

$$\mathcal{P}_{index-shift}(i) = \left(\lambda_{00}^{(N_c-i-1)} \lambda_{01}^{i+1} \mathcal{P}_0 + \lambda_{11}^{i+1} \lambda_{10}^{(N_c-i-1)} \mathcal{P}_1 \right) \binom{N_c}{i+1} \tag{16}$$

Observe that Eq. (16) is usable if there is merely single probable choice for the CH at a particular cluster information (i.e., $N_S = 1$), or the CH counter is infinite (i.e., $\theta_{max} = \infty$). Conversely, when $N_S > 1$ and $\theta_{max} < \infty$, the CH role is disseminated among nodes within a similar set. So, the mean possibility of electing i as the CH for $\theta_{max} < \infty$ is approximately revised by Eq. (17).

$$\bar{\mathcal{P}}_{index-shift}(i) = \frac{\sum_{j=0}^{N_S-1} \mathcal{P}_{index-shift}(j)}{N_S} \tag{17}$$

For $\theta_{max} = 1$, the CH role is rotated uniformly among CMs in a round-robin manner and $\bar{\mathcal{P}}_{index-shift}(i)$ is same for nodes in a similar set. Thus, the CH for each cluster is selected using the PIM scheme, and the ICIC algorithm is used for inter-cluster and intra-cluster communication. The SN receives data from the CHs, makes a global decision on the received data, and determines the DER.

3.5 Calculation of data error rate

The DER quantifies data accuracy as Eq. (18):

$$DER = \mathcal{P}_0\mathcal{P}_F + (1-\mathcal{P}_0)(1-\mathcal{P}_D) \tag{18}$$

In Eq. (18), $\mathcal{P}_F$ and $\mathcal{P}_D$ represent the global false alarm possibility and the global discovery possibility, respectively. To quantify $\mathcal{P}_F$ and $\mathcal{P}_D$, let δ be the amount of 1s in the transferred cluster information and δ' be the amount of 1s in the accepted cluster information. Then, $\mathcal{P}_F$ and $\mathcal{P}_D$ are defined by Eqns. (19) & (20).

$$\mathcal{P}_F = \sum_{\sum_{m=1}^{C}\delta'_m \geq K} \prod_{m=1}^{C} \left(\sum_{i=1}^{N_c} \begin{matrix} Pr.\{\delta_m = i|TA\} \\ Pr.\{\delta'_m|\delta_m = i\} \end{matrix} \right) \tag{19}$$

$$\mathcal{P}_D = \sum_{\sum_{m=1}^{C}\delta'_m \geq K} \prod_{m=1}^{C} \left(\sum_{i=1}^{N_c} \begin{matrix} Pr.\{\delta_m = i|TP\} \\ Pr.\{\delta'_m|\delta_m = i\} \end{matrix} \right) \tag{20}$$

Therefore, $Pr.\{\delta_m = i|TA\}$ and $Pr.\{\delta_m = i|TP\}$ are primarily influenced by the node's local performance and can be defined by Eqns. (21) & (22).

$$Pr.\{\delta_m = i|TA\} = \binom{N_c}{i} \lambda_{01}^{i} \lambda_{00}^{N_c - i} \tag{21}$$

$$Pr.\{\delta_m = i|TP\} = \binom{N_c}{i} \lambda_{11}^{i} \lambda_{10}^{N_c - i} \tag{22}$$

Alternatively, $Pr.\{\delta'_m|\delta_m = i\}$ is primarily determined by Eq. (23).

$$Pr.\{\delta'_m|\delta_m = i\} = \sum_{x^{(p)}\in\chi^{(i)}} \sum_{x^{(q)}\in\chi^{(\delta'_m)}} Pr.\left\{x^{(p)} \to x^{(q)}\right\} \tag{23}$$

In Eq. (23), $\chi^{(i)}$ and $\chi^{(\delta'_m)}$ are sets containing potential cluster information with sums i and δ'_m, respectively. $Pr.\left\{x^{(p)} \to x^{(q)}\right\}$ is the probability of receiving vector $x^{(q)}$ given that the transmitted vector is $x^{(p)}$. The above probability for a given H and $p \neq q$ is represented by the Q function as shown in Eq. (24):

$$Pr.\left\{x^{(p)} \to x^{(q)}\middle|H\right\} = Q\left(\sqrt{\frac{\gamma}{2\sigma^2}}\right), p \neq q \tag{24}$$

In Eq. (24), $\gamma = \|H\Delta\|^2$ and $\Delta = x^{(p)} - x^{(q)}$. Eq. (24) is averaged by adding over the Gamma random factor γ to obtain the closed-form term for $Pr.\left\{x^{(p)} \to x^{(q)}\right\}$ as shown in Eqns. (25) & (26):

$$Pr.\left\{x^{(p)} \to x^{(q)}\right\} = \chi^{N_r} \sum_{t=0}^{N_r - 1} \binom{N_r - 1 + t}{t} (1-\chi)^t \tag{25}$$

Where

$$\chi = \frac{1}{2}\left(1 - \sqrt{\frac{|\Delta|^2}{4\sigma^2 + |\Delta|^2}}\right) \tag{26}$$

In the case of $p = q$, $Pr.\left\{x^{(p)} \to x^{(q)}\right\}$ represented by Eq. (27).

$$Pr.\left\{x^{(p)} \to x^{(q)}\right\}\Big|_{p=q} = 1 - \sum_{x^{(p)}\in\chi^{(i)}} \sum_{x^{(q)}\in\chi^{(\delta'_m)}, q\neq p} Pr.\left\{x^{(p)} \to x^{(q)}\right\} \tag{27}$$

Thus, $\mathcal{P}_F$ and $\mathcal{P}_D$ are calculated by substituting Eqns. (21)-(26) into Eq. (19) or (20), respectively. The resulting formulas are then used in Eq. (18) to obtain the final formula for the mean DER.

Algorithm 1: PIM-ICIC for CH Selection and Energy Minimization in Multi-Hop WSNs

Input: N sensor nodes

Output: ICIC transmission

1. **Begin**
2. Divide N nodes into C cluster with N_c nodes based on proximity;
3. $\boldsymbol{for}(i = 1{:}\,C)$
4. Divide N_c nodes into sets;
5. Choose the CH according to the local cluster data in each cycle;
6. If node extents a threshold limit for being elected as a CH, alternative node from a similar set takes on the CH role;
7. Use the index-shift technique to distribute the CH role evenly among sensor nodes;
8. Transmit the cluster information using the cluster node's indices, where the CH indices are calculated by the total entire cluster information;
9. Calculate the total energy used for transmitting packets according to the total cluster information;
10. Perform ICIC energy minimization approach for inter- and intra-cluster communication;
11. $\boldsymbol{end\ for}$
12. Determine the DER to quantify data accuracy;
13. **End**

4. Simulation results

This subsection provides the performance of the PIM-ICIC and compares it with the existing algorithms such as ICIC [12], HPSO-ILEACH [21] and SCA-Levy [22]. MATLAB 2019b software is used for the simulation analysis. Table 2 presents the parameters and their values utilized to simulate both existing and proposed algorithms, to measure the performance. A comparison is conducted according to the PDR, EC, number of dead nodes and DER.

4.1 PDR

It refers to the number of packets provided to the SN per unit time or round. Fig. 2 compares the PDR of different clustering algorithms for varying numbers of rounds with 400 nodes. The PIM-ICIC algorithm outperforms existing algorithms such as HPSO-ILEACH, SCA-Levy and ICIC by delivering a higher number of packets to the SN.

For instance, PIM-ICIC delivers 5.8×10^5 packets for 1000 rounds, while HPSO-ILEACH, SCA-Levy and ICIC deliver 2.4×10^5, 3.5×10^5 and 4.4×10^5 packets, respectively. This proves that the PIM-ICIC increases the overall quantity of transferred data to the SN compared to the others by selecting CHs equally among nodes and reducing the DER.

4.2 Network energy consumption

It is the proportion of the energy consumed by the nodes to the overall initial energy accessible during node deployment. It is calculated by Eq. (28).

$$Network\ EC = \frac{\sum_{i=1}^{M} E_{used_i}}{\sum_{i=1}^{N} E_{initial_i}} \quad (28)$$

In Eq. (28), E_{used_i} is the energy used by node i during its deployment, and $E_{initial\_i}$ is the initial energy available before deployment for i. Fig. 3 compares the PDR for different clustering algorithms with varying numbers of nodes over 1000 rounds.

The PIM-ICIC outperforms HPSO-ILEACH, SCA-Levy and ICIC by delivering a higher number of packets to the SN. For example, PIM-ICIC delivers 56.82%, 46.81%, and 23.21% more packets to the SN for 400 nodes in 1000 rounds compared to HPSO-ILEACH, SCA-Levy and ICIC, respectively.

Fig. 4 compares network EC of various clustering algorithms during node deployment with different numbers of nodes over 1000 rounds. It can be observed that the PIM-ICIC is more efficient in total network EC for 400 nodes due to its uniform EC mechanism, reducing EC by 15.22%, 13.33%, and

Table 2. Simulation parameters

Parameters	Value
Simulation region	1000×1000 m^2
No. of sensor nodes	400
No. of clusters	20
No. of nodes per clusters	20
No. of sets	20
No. of nodes per set	20
SN position	(420, 400)
Area of each cluster	100×100 m^2
Network topology	Flat grid
Antenna	Omni antenna
No. of receive antennas (N_r)	5
Channel	Wireless
Propagation category	Two ray ground
MAC layer	IEEE802.11
MAC protocol	TDMA
Packet size	500 bytes
Traffic source	Constant Bit Rate (CBR)
Initial energy of nodes	5 mJ
Threshold energy	0.1 mJ
ϵ_{fs}	10 pJ/bit/m^2
ϵ_{mp}	0.0013 pJ/bit/m^4
E_{elec}	50 nJ/bit
Energy used in data collection	5 nJ/bit
No. of rounds	1000
Transmission range	50 m
Simulation time	150 sec
Total absence probability ($\mathcal{P}_0$)	0.75
Total presence probability ($\mathcal{P}_1 = 1 - \mathcal{P}_0$)	0.75
False-alarm probability (λ_{01})	0.3
Miss-detection probability (λ_{10})	0.3
True detection probability for absent (λ_{00})	0.7
True detection probability for present (λ_{11})	0.7
Maximum CH counter	1000
Maximum transmission number	1000
SNR	0-45dB

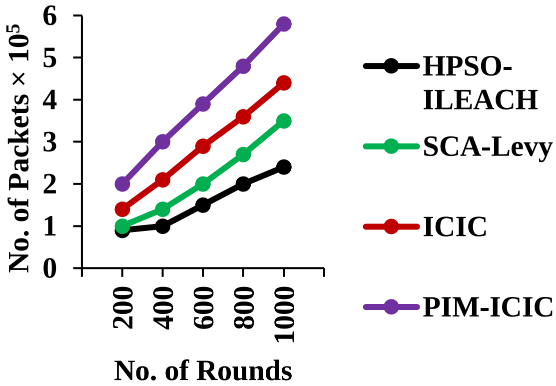

Figure. 2 PDR vs. rounds for 400 nodes

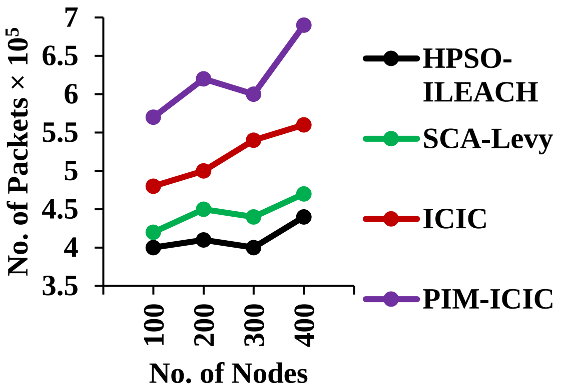

Figure. 3 PDR vs. No. of nodes for 1000 rounds

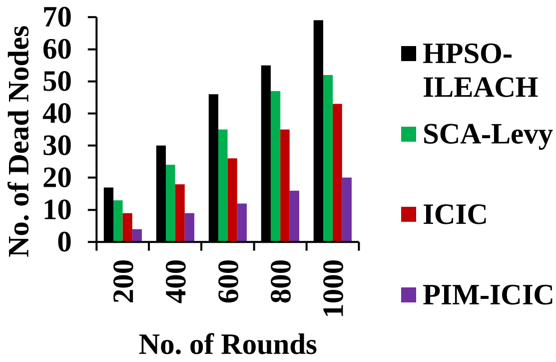

Figure. 6 No. of dead nodes vs. rounds for 400 nodes

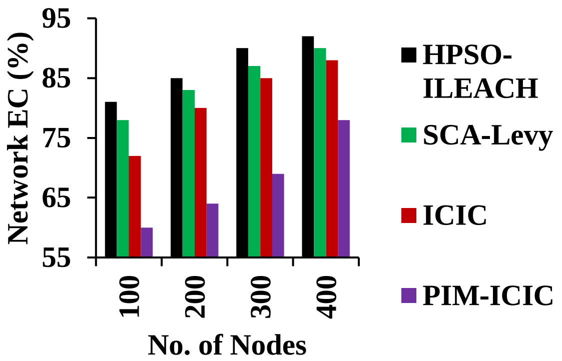

Figure. 4 Network EC vs. No. of nodes for 1000 rounds

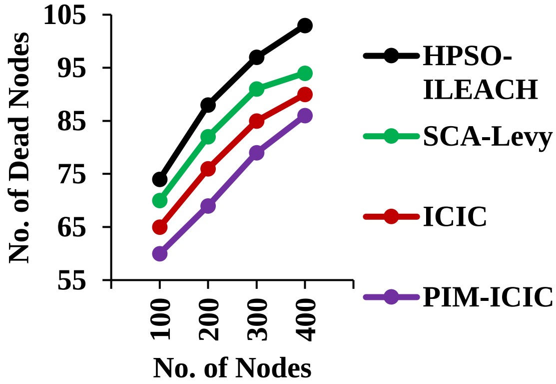

Figure. 7 No. of dead nodes vs. No. of nodes for 1000 rounds

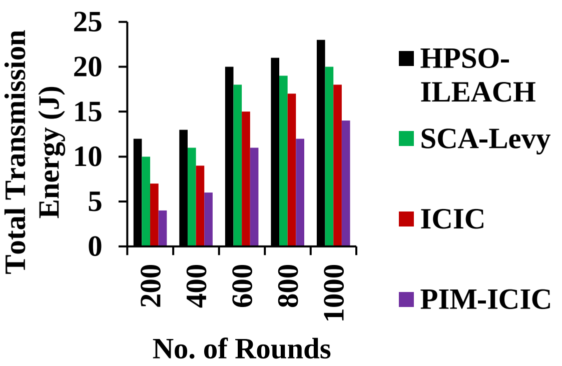

Figure. 5 Network EC in data transfer vs. rounds for 400 nodes

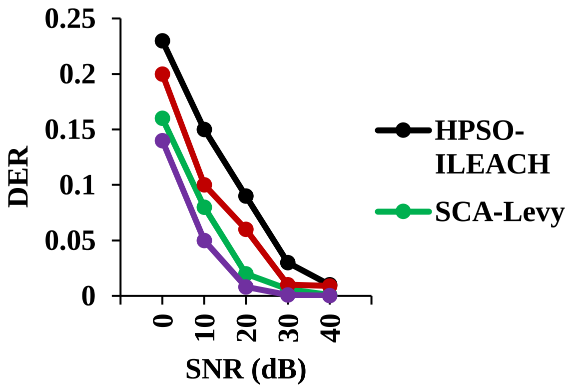

Figure. 8 DER vs. SNR

11.36% compared to HPSO-ILEACH, SCA-Levy and ICIC algorithms, respectively.

4.3 Network energy utilization in data transfer

It is the energy consumed by all nodes in receiving data from nearby CH node and transferring it to the SN. It is calculated by Eq. (29).

$$E_{total} = \sum_{i=1}^{M} E_{totaltrans\_i} + E_{totalRec\_i} \quad (29)$$

In Eq. (29), M is the maximum amount of nodes, $E_{totaltrans\_i}$ is the overall transfer energy used by i, and $E_{totalRec\_i}$ is the total energy used by i while receiving data from nearby node.

Fig. 5 compares the network EC of different clustering algorithms for data transmission with varying rounds for 400 nodes. PIM-ICIC reduces total transmission energy by 39.13%, 30%, and 22.22% compared to HPSO-ILEACH, SCA-Levy and ICIC, respectively for 1000 rounds. PIM-ICIC saves energy during data transmission by efficiently assigning the CH role among nodes.

4.4 Number of dead nodes

It is the overall amount of nodes that have depleted their energy and are unable to function as node of WSN. Fig. 6 shows the comparison of dead nodes in various clustering algorithms over different rounds with 400 nodes. The PIM-ICIC algorithm reduces dead nodes by 71.01%, 61.54%, and 53.49% compared to HPSO-ILEACH, SCA-Levy and ICIC, respectively for 1000 rounds. PIM-ICIC also achieves the longest network longevity by effectively distributing the CH role among sensor nodes in multi-hop WSNs, leading to improved PDR and energy efficiency.

Fig. 7 compares dead nodes in different clustering algorithms with varying numbers of nodes over 1000 rounds. The PIM-ICIC algorithm reduces dead nodes by 16.5%, 8.51%, and 4.44% compared to HPSO-ILEACH, SCA-Levy and ICIC, respectively when deploying 400 nodes. This suggests that the PIM-ICIC algorithm offers a longer network lifetime in multi-hop WSNs.

4.5 DER

The mean DER is calculated across SNR values ranging between 0-45dB. SNR is the relation of transmit power to noise power, with the transmit power consistently set to unity. Fig. 8 compares DER for different clustering algorithms at various SNRs. The PIM-ICIC algorithm outperforms HPSO-ILEACH, SCA-Levy and ICIC by reducing DER by 98%, 77.78%, and 97.78%, respectively, at 40 dB SNR. This indicates that the PIM-ICIC algorithm can significantly improve data reception accuracy compared to other algorithms, striking a balance between network lifetime and DER.

5. Conclusion

This paper introduces a novel PIM-ICIC algorithm for multi-hop WSNs. It evenly distributes the energy-intensive CH functionality among nodes using an index-shift scheme to lengthen the network lifespan. CMs are organized into sets that coordinate CH selection based on local cluster information and previous CH counts. Simulation results demonstrated that the PIM-ICIC outperforms current clustering schemes in terms of PDR, energy efficiency, reduced dead nodes, and DER across various scenarios. By balancing network longevity and data accuracy, PIM-ICIC shows promise for improving multi-hop WSN performance. The results show that the PIM-ICIC system achieved a PDR of 5.8×10^5 packets, consumed 14J of energy, and had 20 dead nodes over 1000 rounds. With 400 nodes in the network, the PIM-ICIC system achieved a PDR of 6.9×10^5 packets, consumed 78% of the energy, and had 86 dead nodes. Also, at an SNR of 40dB, the system achieved a DER of 0.0002, outperforming existing algorithms such as ICIC, HPSO-ILEACH, SCA-Levy. However, the network lifespan decreases as the number of nodes increases. Therefore, future work will focus on developing relay node selection based on various criteria such as the number of nodes closer to the BS, coverage range and other relevant parameters to improve network longevity.

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

Conceptualization, methodology, software, validation, Monika; formal analysis, investigation, Vijayalakshmi; resources, data curation, writing—original draft preparation, Monika; writing—review and editing, Monika; visualization, supervision, Vijayalakshmi;

References

[1] N. Hossain, M. I. H. Rimon, M. A. Mimona, M. H. Mobarak, J. Ghosh, M. A. Islam, and M. Z. Al Mahmud, "Prospects and Challenges of Sensor Materials: A Comprehensive Review", *e-Prime-Advances in Electrical Engineering, Electronics and Energy*, Vol. 7, p. 100496, 2024.

[2] A. Rady, E. S. M. El-Rabaie, M. Shokair, and N. Abdel-Salam, "Comprehensive Survey of Routing Protocols for Mobile Wireless Sensor Networks", *International Journal of Communication Systems*, Vol. 34, No. 15, p. e4942, 2021.

[3] C. Nakas, D. Kandris, and G. Visvardis, "Energy Efficient Routing in Wireless Sensor Networks: A Comprehensive Survey", *Algorithms*, Vol. 13, No. 3, p. 72, 2020.

[4] M. Shafiq, H. Ashraf, A. Ullah, and S. Tahira, "Systematic Literature Review on Energy Efficient Routing Schemes in WSN–A Survey", *Mobile Networks and Applications*, Vol. 25, pp. 882-895, 2020.

[5] P. Rawat and S. Chauhan, "Clustering Protocols in Wireless Sensor Network: A Survey, Classification, Issues, and Future Directions", *Computer Science Review*, Vol. 40, p. 100396, 2021.

[6] I. Daanoune, B. Abdennaceur, and A. Ballouk, "A Comprehensive Survey on LEACH-Based Clustering Routing Protocols in Wireless Sensor

Networks", *Ad Hoc Networks*, Vol. 114, p. 102409, 2021.

[7] R. Kavra, A. Gupta, and S. Kansal, "Systematic Study of Topology Control Methods and Routing Techniques in Wireless Sensor Networks", *Peer-to-Peer Networking and Applications*, Vol. 15, No. 4, pp. 1862-1922, 2022.

[8] M. K. Khan, M. Shiraz, Q. Shaheen, S. A. Butt, R. Akhtar, M. A. Khan, and W. Changda, "Hierarchical Routing Protocols for Wireless Sensor Networks: Functional and Performance Analysis", *Journal of Sensors*, Vol. 2021, pp. 1-18, 2021.

[9] W. Kim, M. M. Umar, S. Khan, and M. A. Khan, "Novel Scoring for Energy-Efficient Routing in Multi-Sensored Networks", *Sensors*, Vol. 22, No. 4, p. 1673, 2022.

[10] A. M. Jubair, R. Hassan, A. H. M. Aman, H. Sallehudin, Z. G. Al-Mekhlafi, B. A. Mohammed, and M. S. Alsaffar, "Optimization of Clustering in Wireless Sensor Networks: Techniques and Protocols", *Applied Sciences*, Vol. 11, No. 23, p. 11448, 2021.

[11] A. J. Williams, M. F. Torquato, I. M. Cameron, A. A. Fahmy, and J. Sienz, "Survey of Energy Harvesting Technologies for Wireless Sensor Networks", *IEEE Access*, Vol. 9, pp. 77493-77510, 2021.

[12] I. K. Shah, T. Maity, Y. S. Dohare, D. Tyagi, D. Rathore, and D. S. Yadav, "ICIC: A Dual Mode Intra-Cluster and Inter-Cluster Energy Minimization Approach for Multihop WSN", *IEEE Access*, Vol. 10, pp. 70581-70594, 2022.

[13] A. S. Sadeq, R. Hassan, H. Sallehudin, A. H. M. Aman, and A. H. Ibrahim, "Conceptual Framework for Future WSN-MAC Protocol to Achieve Energy Consumption Enhancement", *Sensors*, Vol. 22, No. 6, p. 2129, 2022.

[14] S. Althunibat and R. Mesleh, "Index Modulation for Cluster-Based Wireless Sensor Networks", *IEEE Transactions on Vehicular Technology*, Vol. 67, No. 8, pp. 6943-6950, 2018.

[15] S. E. Pour and R. Javidan, "A New Energy Aware Cluster Head Selection for LEACH in Wireless Sensor Networks", *IET Wireless Sensor Systems*, Vol. 11, No. 1, pp. 45-53, 2021.

[16] M. Singh and S. K. Soni, "Network Lifetime Enhancement of WSNs Using Correlation Model and Node Selection Algorithm", *Ad Hoc Networks*, Vol. 114, p. 102441, 2021.

[17] 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, 2021.

[18] K. K. Le-Ngoc, Q. T. Tho, T. H. Bui, A. M. Rahmani, and M. Hosseinzadeh, "Optimized Fuzzy Clustering in Wireless Sensor Networks Using Improved Squirrel Search Algorithm", *Fuzzy Sets and Systems*, Vol. 438, pp. 121-147, 2022.

[19] M. Rami Reddy, M. L. Ravi Chandra, P. Venkatramana, and R. Dilli, "Energy-Efficient Cluster Head Selection in Wireless Sensor Networks Using an Improved Grey Wolf Optimization Algorithm", *Computers*, Vol. 12, No. 2, p. 35, 2023.

[20] M. Tay and A. Senturk, "A New Energy-Aware Cluster Head Selection Algorithm for Wireless Sensor Networks", *Wireless Personal Communications*, Vol. 122, No. 3, pp. 2235-2251, 2022.

[21] S. Sharmin, I. Ahmedy, and R. Md Noor, "An Energy-Efficient Data Aggregation Clustering Algorithm for Wireless Sensor Networks Using Hybrid PSO", *Energies*, Vol. 16, No. 5, pp. 1-24, 2023.

[22] X. Guo, Y. Ye, L. Li, R. Wu, and X. Sun, "WSN Clustering Routing Algorithm Combining Sine Cosine Algorithm and Lévy Mutation", *IEEE Access*, Vol. 11, pp. 22654-22663, 2023.

[23] I. Kumar Shah, T. Maity, and Y. Singh Dohare, "Algorithm for Energy Consumption Minimisation in Wireless Sensor Network", *IET Communications*, Vol. 14, No. 8, pp. 1301-1310, 2020.