Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A Lightweight, Distributed Energy-Aware Clustering Algorithm for Heterogeneous IoT Sensor Networks Using Adaptive Thresholding

Naeem A. Askar
University of Duhok
Ismail Y. Maolood
Sulaimani Polytechnic University
Azad A. Ameen
azad.ameen@chu.edu.iq

Charmo University

Research Article

Keywords: IoT, Clustering, Heterogeneity, Energy Efficiency, Adaptive Thresholding

Posted Date: September 14th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10968306/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

A Lightweight, Distributed Energy-Aware Clustering Algorithm for Heterogeneous IoT Sensor Networks Using Adaptive Thresholding

Naeem A. Askar [1], Ismail Y. Maolood [2,3], Azad A. Ameen [4]

[1]Department of Mathematics, College of Basic Education, University of Duhok, Kurdistan Region – F.R. Iraq

[2]Department of Information Technology, Technical College of Informatics, Sulaimani Polytechnic University, Sulaimani, Kurdistan Region– F.R. Iraq

[3]Directorate of Statistics and Planning, General Directorate of Study, Planning and Follow-Up, Ministry of Higher Education and Scientific Research, Kurdistan

1st Author's organizational email: naeem.askar@uod.ac
1st Author's ORCID link: https:// https://orcid.org/0009-0007-5882-1919
1st Author's Google Scholar Citation link: https:// https://scholar.google.com/citations?user=UTSEvxYAAAAJ&hl=en&authuser=1

2nd Author's organizational email: ismail.maulood@mhe-krg.org
2nd Author's ORCID link: https:// https://orcid.org/0000-0003-1683-1493
2nd Author's Google Scholar Citation link: https:// https://scholar.google.com/citations?user=tdB8KHQAAAAJ&hl=en

4nd Author's organizational email: azad.ameen@chu.edu.iq
4nd Author's ORCID link: https:// https://orcid.org/0000-0001-9350-2330
4nd Author's Google Scholar Citation link: https:// https://scholar.google.com/citations?user=4PZE4eUAAAAJ&hl=en

ABSTRACT

Heterogeneous IoT sensor networks, characterized by nodes with diverse energy capacities, computational capabilities, and communication ranges, introduces significant challenges for maximizing network lifetime and maintaining efficient network operation. Existing clustering algorithms depend on centralized control, fixed thresholds values, or assumptions of homogeneous node characteristics, which limitations may reduce their effectiveness in heterogeneous IoT environments. To overcome these limitations, the study proposes AdaHet-Clustm a distributed clustering algorithm that dynamically chooses the cluster head heads based on local remaining energy conditions and adapts its operation according the network's energy consumption patterns. The proposed algorithm is supported by a rigorous theoretical framework comprising seven formal theorems that analyze energy balancing, convergence properties, scalability and computational complexity. Extensive simulation results demonstrate that AdaHet-Clut significantly enhances network performance, achieving improvements in First Node Death (FND) of 16.3% and 49% compared with HMGEAR and DEEC, respectively. Moreover, the proposed algorithm reduces control overhead by 24.1% relative to HMGEAR while increasing overall energy efficiency and extend network lifetime. The results demonstrate that adaptive thresholding can significantly enhance clustering performance in heterogeneous IoT sensor networks while maintaining low communication overhead. This work informs future developments in large-scale, heterogeneous IoT deployments where node diversity, resource constraints, and network longevity are critical design considerations.

Index Terms— IoT, Clustering, Heterogeneity, Energy Efficiency, Adaptive Thresholding.

I. INTRODUCTION

The Internet of Things (IoT) has revolutionized data driven applications across environmental monitoring, industrial automation, and smart cities (Mohammed et al., 2022, Ekler et al., 2022, Askar, 2023). At the core of IoT ecosystems lie Wireless Sensor Networks (WSNs) comprising resource-constrained nodes that collect and transmit environmental data(Al-Fuqaha et al., 2015, Atzori et al., 2010, Askar et al., 2023). However, IoT sensor nodes are battery-powered and deployed in remote locations, making energy efficiency the most critical challenge for network longevity (Reina et al., 2013, Zhu et al., 2015).

Clustering is one of the most prominent mechanisms for achieving energy efficiency in WSNs (Farooq and Abdul, 2018, Lonkar et al., 2026). Using intra-cluster and inter-cluster communication, cluster-based architectures significantly reduce data transmissions(Alomari et al., 2022, Katiyar et al., 2011). However, clustering effectiveness depends critically on cluster head (CH) selection, as CHs bear the disproportionate energy burden of data aggregation and long-distance transmission(Baranidharan and Santhi, 2015). Poor CH selection leads to premature node death, network fragmentation, and reduced coverage(Sharma et al., 2026, Ordouei, Dwivedi and Sharma, 2020).

In homogeneous networks, randomized CH rotation mechanisms such as LEACH (Heinzelman et al., 2000) achieve reasonable energy distribution. However, in heterogeneous IoT sensor networks where nodes vary in battery capacity, processing power, and communication range simple randomized or static weighting approaches prove inadequate. Low-capacity CHs deplete rapidly, creating energy holes, while consistently selecting only high-capacity nodes concentrates energy drain on a subset(Goyal et al., 2021, Moasses et al., 2021).

Existing clustering algorithms exhibit several inherent limitations when deployed in heterogeneous internet of things environments. Traditional protocols such as LEACH and HEED are designed under the assumption of homogeneous nodes, making them unsuitable for networks with diverse energy capacities. DEEC and EDEEC use static weights that cannot adapt to dynamic changes in residual energy. Similarly, EA-CRP depend on semi-centralized coordination, which have high communication overhead and less scalability. HMGEAR requires dedicated gateway infrastructure, introducing additional hardware cost while creating a potential single point of failure(Al-Hussain and Al-Suhail, 2025). Machine (Al-Hussain and Al-Suhail, 2025). Furthermore, machine learning approaches like HEOCP require large training datasets and GPU acceleration. Metaheuristic algorithms like CLLDOA-ECP operate centrally with limited scalability and more computational complexity in large network (Dwivedi and Sharma, 2020). Thus, a clustering system is needed that effectively handles node heterogeneity without relying on static thresholds, centralized coordination, or external infrastructure. The system must adapt continuously to energy depletion patterns, resolve CH contention deterministically, and operate within resource constraints (Haris and Nam, 2025, Alsharari et al., 2020).

Recent studies have explored alternative solutions for improving energy efficiency in IoT sensor networks. Optimization-based approaches, such as the Enhanced Particle Swarm Optimization (EPSO) method proposed by(Lilo et al., 2024) have demonstrated significant reductions in transmission energy consumption and improvements in network lifetime. However, such methods generally require iterative optimization procedures and additional computational resources, which may not be suitable for resource-

constrained IoT devices. Similarly, recent research on mitigating the impact of dead nodes has emphasized the importance of balanced energy consumption and prolonged node survivability for maintaining network connectivity and coverage in WSNs. These findings highlight the need for lightweight and adaptive energy management mechanisms capable of operating efficiently in heterogeneous IoT environments (Khorsheed, 2025).

To our knowledge, the combination of dynamic local energy ratios, real-time adaptive thresholding, and deterministic two-hop collision resolution has not been addressed for fully distributed heterogeneous clustering. Hence, this paper introduces AdaHet-Clust (Adaptive Heterogeneous Clustering). This proposed work based on an adaptive thresholding mechanism that responds to overall network energy consumption, a dynamic heterogeneity factor that prioritizes nodes with higher energy capacity, and a deterministic scoring function for collision-free CH placement. Accordingly, the main contributions of this study are:

- We propose a lightweight, fully distributed clustering algorithm that enables each sensor node to independently determine its cluster head (CH) election probability using only local available information.
- We incorporate a real-time adaptive thresholding mechanism that dynamically adjusts the cluster head election process in responds to network energy depletion without manual reconfiguration.
- We developed a deterministic two-hop collision resolution mechanism ensuring that among competing CHs within two-hop distance, the node with the highest residual energy and lowest node degree is selected.
- Extensive performance evaluation comparing AdaHet-Clust with state-of-the-art protocols (LEACH, DEEC, HMGEAR), demonstrating improvements in network lifetime, energy uniformity, and control overhead.

The remainder of this paper is organized as follows. Section II provides a comprehensive literature review. Section III describes the system model and proposed algorithm. Section IV. Describe the theoretical analysis. Section V. describes simulation results and discusses. Section VI concludes and future work the paper.

II. Literature Review

Clustering has been extensively studied as an energy-efficient data collection mechanism for wireless sensor networks (WSNs) and IoT environments. This section provides a comprehensive review of

existing clustering algorithms, categorizing them by their fundamental approaches, and identifies the research gap that motivates our proposed AdaHet-Clust algorithm, that is summarized in Table I.

The Authors in(Heinzelman et al., 2000) introduced LEACH (Low-Energy Adaptive Clustering Hierarchy), which remains the most cited clustering protocol in WSN literature. LEACH employs randomized rotation of cluster heads (CHs) to distribute energy consumption evenly across all nodes. Each node independently generates a random number between 0 and 1 and compares it against a dynamic threshold T(n)=P/(1−P·(r*mod1/P)), where P is the desired CH probability and r is the current round number. While LEACH is fully distributed and requires no training data, it makes a critical assumption of node homogeneity all nodes are assumed to have identical initial energy, processing capability, and communication range. This assumption fails in real-world heterogeneous IoT deployments, leading to suboptimal performance where low-energy nodes are selected as CHs and die prematurely.

In(Ossama, 2004) the authors suggested the Hybrid Energy-Efficient Distributed clustering (HEED) clustering protocol as an enhancement to LEACH by incorporating both residual energy and intra-cluster communication cost into the cluster head (CH) selection process. HEED employs an iterative CH election mechanism in which each node announces its candidacy with probability proportional to their remaining energy. The clustering process terminates after a fixed number of iterations, ensuring adequate CH coverage. Compared with LEACH, HEED achieves improved energy balancing and introduces a basic collision resolution strategy that prioritizes nodes with higher remaining energy during CH selection. However, the protocol retains the homogeneous node assumption and requires multiple control message exchanges during each cluster round, resulting increasing communication overhead and energy consumption.

In (Nehra and Sharma, 2013) the researchers proposed Power-Efficient Gathering in Sensor Information Systems named PEGASIS, which forms chains of nodes rather than clusters. Each node communicates only with its nearest neighbor, and a designated leader node aggregates data and transmits to the base station. While PEGASIS decrease transmission distance, it introduces significant latency and lacks fault tolerance, as the failure of any node breaks the chain. In(Qing et al., 2006), the researcher introduces the DEEC (Distributed Energy-Efficient Clustering), one of the first algorithms explicitly designed for heterogeneous WSNs. DEEC uses static weights based on initial energy ratios—advanced nodes with higher initial energy receive proportionally higher CH probabilities. The CH probability for node ii is defined as $P_{opt} * E_i(r)/E'(r)$, where $E_i(r)$is residual energy and $E'(r)$is average network energy. While DEEC improves network lifetime in two-level heterogeneity scenarios, its weights are

predetermined and cannot adapt to dynamic changes in residual energy distribution.

In(Javaid et al., 2013) the researchers introduced the EDEEC (Enhanced Distributed Energy-Efficient Clustering), which extends DEEC to handle three-level heterogeneity by introducing a "super node" category with the highest initial energy. EDEEC improves upon DEEC but inherits its limitations static thresholds and lack of real-time adaptation. Similarly, DDEEC (Developed DEEC) introduced threshold-based CH selection but still relies on round-based probability decay.

EA-CRP (Energy-Aware Clustering Routing Protocol) are proposed by(Darabkh et al., 2018) as an energy-aware approach for heterogeneous IoT networks. EA-CRP incorporates energy consumption prediction into the CH selection process, using a weighted metric that considers residual energy, distance to base station, and node degree. The protocol operates in a semi-distributed manner, where CH candidates are evaluated based on a composite score. EA-CRP achieves improved network lifetime compared to LEACH and DEEC but has several limitations: (1) it requires periodic global information collection at the base station; (2) it uses static weights that do not adapt to network dynamics; (3) it lacks a formal collision resolution mechanism; and (4) its performance degrades as network size increases due to centralized coordination overhead. Despite being published in 2018, EA-CRP remains a relevant baseline for energy-aware clustering in heterogeneous environments.

In (Ibriq, 2004) the authors employed a genetic algorithms (GA) to optimize the cluster head (CH) selection process, thereby improving energy efficiency and extend network lifetime. The genetic algorithm optimization searches for near-best CH configuration, but it depends on centralized computation and require complete knowledge of network wide information. Moreover, the iterative optimization process introduces considerable computational complexity and execution overhead limiting its suitability for IoT sensor networks.

The Authors in (Ti et al., 2026) introduced HEOCP (Heterogeneous Energy-Efficient Optimized Clustering Protocol), which uses a ResNet-50 deep neural network for CH prediction. HEOCP achieves high accuracy in CH selection by training on labeled network state data. However, it requires extensive training datasets, GPU acceleration for inference, and cloud or edge infrastructure for deployment. These requirements make HEOCP impractical for low-power IoT nodes with limited memory and processing capability.

The authors in (Jibreel et al., 2022) proposed HMGEAR (Heterogeneous Gateway-based Energy-Aware Multi-hop Routing Protocol) as an enhancement to the homogeneous MGEAR protocol. HMGEAR divides the sensing field into four logical regions based on predefined distance thresholds. A gateway node is placed at the network center while the base station is positioned outside the sensing area.

Homogeneous nodes in regions 1 and 2 transmit data directly to the base station or gateway, respectively. Heterogeneous nodes with higher energy are deployed in regions 3 and 4, where cluster heads are selected based on residual energy using the probabilities A multi-hop communication strategy is employed from regions 3 and 4 to the gateway, reducing energy depletion of distant nodes. An energy hole elimination technique prevents nodes from transmitting when their energy falls below a threshold. It requires dedicated gateway infrastructure and uses static CH selection probabilities based on initial energy ratios HMGEAR cannot dynamically adapt to real-time energy distribution changes and relies on gateway coordination, creating a single point of failure and additional hardware costs.

The researcher in (Sun et al., 2026) proposed CLLDOA-ECP (Chaotic Lévy Lens Dream Optimization Algorithm for clustering), a metaheuristic approach specifically designed for Industrial IoT (IIoT) environments. The algorithm extends network lifetime by 36.6% compared to existing benchmarks. CLLDOA-ECP uses a chaotic Lévy flight mechanism to explore the solution space and a lens dreaming strategy to escape local optima. However, the algorithm operates in a centralized manner, requiring a powerful central controller to execute the optimization. The computational complexity of metaheuristic algorithms scales poorly with network size, making CLLDOA-ECP unsuitable for large-scale or resource-constrained deployments.

In (Ab Rahman and Hamim, 2025) the Authors proposed EECR (Improved K-means clustering with adaptive distance threshold). This protocol combines an improved K-means algorithm for initial cluster formation with an adaptive distance threshold that adjusts based on node proximity. EECR reduces energy consumption by 5% and extends network lifetime by 84% compared to HEED. While EECR employs an adaptive threshold mechanism, the adaptation is based on distance rather than network-wide energy consumption feedback. Consequently, EECR cannot respond to energy depletion dynamics nodes may die prematurely while distance thresholds remain unchanged.

TABLE I
COMPREHENSIVE COMPARISON OF EXISTING CLUSTERING ALGORITHMS

Algorithm	Year	Distributed	Heterogeneity Type	Adaptive Threshold	Collision Resolution	Gateway Required
LEACH (Heinzelman et al., 2000)	2000	✓	None	✗	✗ (random)	✗
HEED(Ossama, 2004)	2004	✓	None	✗	✓ (residual)	✗
PEGASIS(Nehra and Sharma, 2013)	2002	✓	None	✗	✗	✗
EDEEC(Javaid et al., 2013)	2009	✓	Multi-level (static)	✗	✗	✗
EA-CRP (Darabkh et al., 2018)	2018	✓	static	✗	✗	✗
GA-Clustering (Ibriq, 2004)	2006	✗	Multi-level	✗	✗	✗
HEOCP(Ti et al., 2026)	2024	✗	Static	✗	✗	✗
HMGEAR (Jibreel et al., 2022)	2022	✗ (gateway)	Static (2 – level)	✗ (distance)	✗	✓
CLLDOA-ECP (Sun et al., 2026)	2026	✗	Dynamic	✗	✗	✗
EECR (Ab Rahman and Hamim, 2025)	2025	✓	Static	✓ distance	✗	✗
AdaHet-Clust	2026	✓	Dynamic (α factor)	✓ (energy)	✓ (2-hop)	✗

The literature review demonstrates that existing clustering algorithms suffer from fundamental limitations. LEACH and HEED assume node homogeneity. DEEC and EDEEC use static weights that cannot adapt to energy dynamics. EA-CRP relies on semi-centralized coordination and static weights.

HMGEAR requires a dedicated gateway, creating hardware cost and a single point of failure. HEOCP and CLLDOA-ECP demand datasets and GPU acceleration. EECR uses distance-based adaptation but cannot respond to energy depletion. This gap motivates AdaHet-Clust, which achieves fully distributed operation, dynamic heterogeneity handling, real-time energy-based adaptive thresholding, deterministic collision resolution, and zero dependency on datasets, training, or gateway infrastructure.

III. PROPOSED ALGORITHM: ADAHET-CLUST

A. System Model

We consider a network of N heterogeneous IoT sensor nodes deployed uniformly at random in a 200 ×200 m² area. There are three levels of (Medium (1.0 J), and High (2.0 J)). Initial energy $E_{init}(i) = E_{max}(i)$ the base station (BS) is Location at (100,250), outside the deployment area. Nodes are location-aware and can adjust transmission power. The first-order radio model(Heinzelman et al., 2000) [21] defines energy consumption for transmission and reception. Nodes are location-aware (GPS or localization); each node knows its residual energy; network is static or low-mobility.

B. Energy Consumption Model

We adopt the first-order radio model (Heinzelman et al., 2000). To transmit a *k*-bit message over distance *d*, the energy consumed is:

$$E_{tx}(k,d) = \begin{cases} k * E_{elec} + k * \epsilon_{fs} * d^2, & d < d_0 \\ k * E_{elec} + k * \epsilon_{np} * d^4, & d \geq d_0 \end{cases} \quad (1)$$

To receive a *k*-bit message, the energy consumed is:

$$E_{rx}(k) = k * E_{elec} \quad (2)$$

Where $d_0 = \sqrt{\epsilon_{fs}/\epsilon_{mp}}$ is the threshold distance.

C. Algorithm Overview

AdaHet-Clust operates in rounds, each consisting of five phases: Setup, CH Election, Collision Resolution, Cluster Formation, and Data Transmission. Fig. 1. presents the flowchart of the proposed AdaHet-Clust algorithm, illustrating the five sequential phases of operation

D. Phase 1: Setup

At the beginning of each round (or periodically every K rounds to reduce overhead), each node broadcasts its current residual energy $E_{res}(i)$ to all one hop neighbors. Each node calculates the local

average energy, after collecting energy information from neighbors:

$$E_{avg_local}(i) = \frac{E_{res}(i)+\sum_{j\in N_1(i)} E_{res}(j)}{1+|N\prime_1(i)|} \quad (3)$$

where $N_1(i)$ is the set of one hop neighbors of node *i*, and $N'_1(i)$ is its degree

E. Phase 2: CH Election

Every node *i* independently computes its CH election probability:

$$P_{ch}(i) = P_{opt} x \frac{E_{res}(i)}{E_{avg_local}} * \alpha(i) * \beta(i) * \gamma(r) \quad (4)$$

Where: $$\alpha(i) = \frac{E_{max}(i)}{max_{j\in N_1\cup\{i\}} E_{max(j)}} \quad (5)$$

$$\beta(i) = 1 - \frac{d(i,BS)}{d_{max}}, \quad d_{max} = max_{k\in N} d(k, BS) \quad (6)$$

$$\gamma(r) = \frac{N_{alive}(r)}{N} \quad (7)$$

The adaptive threshold is:

$$T_{adapt}(r) = T_{base} * \frac{E_{total_res}(r)}{E_{total_init}} \quad (8)$$

Where:

$$T_{base} = 0.1 * P_{opt} = 0.005, E_{total_init} = \sum_i E_{init}(i), and\ E_{total_res}(r) is\ estimated\ via\ periodic\ CH\ to\ BS\ reports.$$

$$Node\ i\ broadcasts\ its\ CH\ candidacy\ only\ if\ p_{ch}(i) > T_{adapt}(r) \text{ and random } () < p_{ch}(i) \quad (9)$$

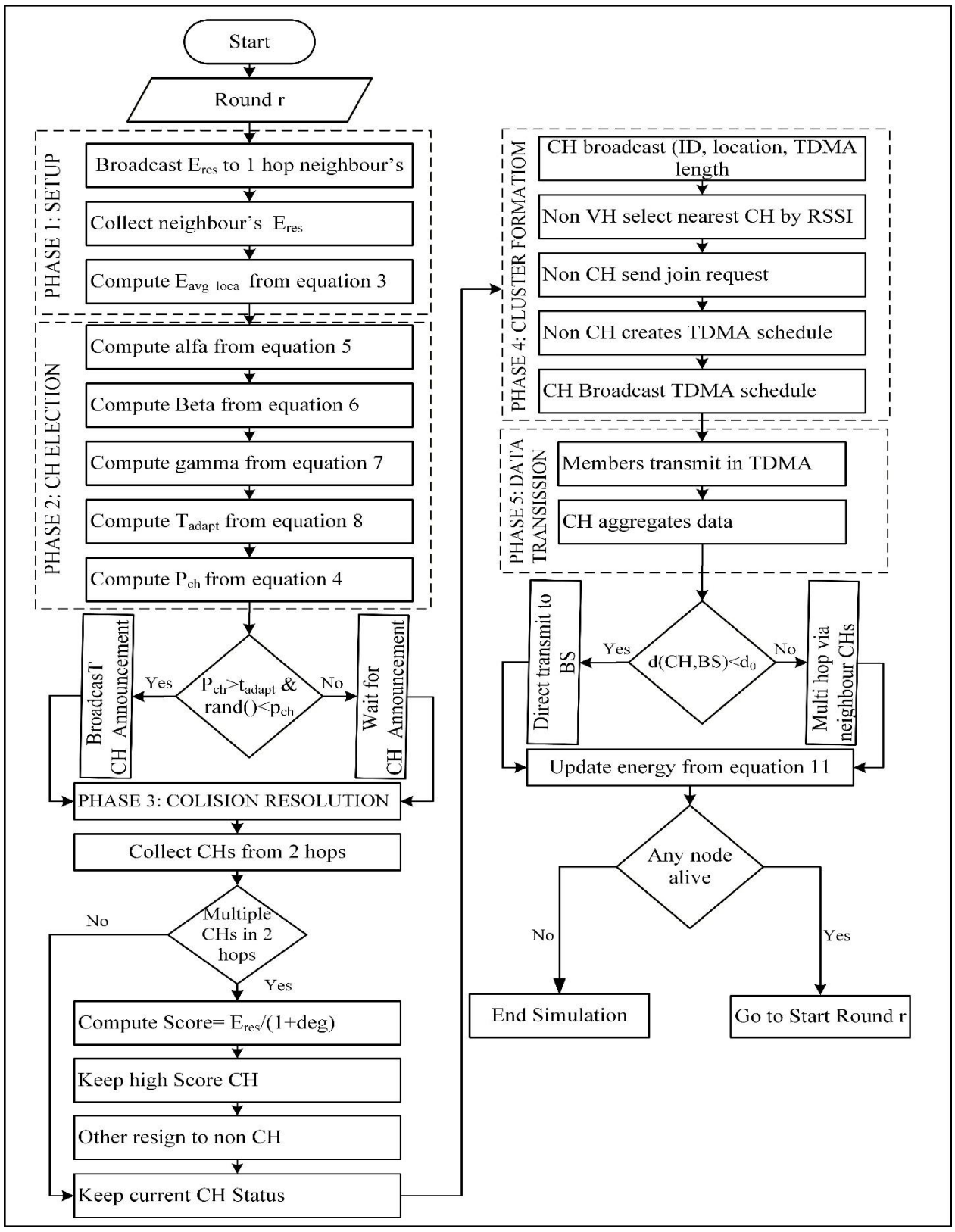


Fig 1: the Flowchart of AdaHet-Clust algorithm.

F. Phase 3: Collision Resolution

If multiple CHs are detected within 2-hop distance, deterministic resolution is triggered. Each competing CH:

$$core(i) = E_{res}(i) * \frac{1}{1 + \deg(i)} \quad (10)$$

where $\deg(i)$ is the number of 1-hop neighbors. Among CHs within 2 hops, the node with the highest Score remains CH. Ties broken by *Emax(i)* then lower ID.

G. Phase 4: Cluster Formation

- Each CH broadcasts an advertisement message containing (CH ID, position, TDMA frame length).
- Each non-CH nodes select the nearest CH based on received signal strength (RSSI).
- Each non-CH sends join request to its selected CH.
- Each CH creates a TDMA schedule for its members.

H. Phase 5: Data Transmission

This phase handles actual data transmission and energy consumption.

For Cluster Members (Non-CH):

- Wait for assigned TDMA slot
- Transmit sensed data to CH

For Cluster Heads (CH):

- Receive data from members in TDMA order
- Aggregate received data (e.g., average, max, min)
- The CH forwards aggregated data to BS:
    - $if\ d(CH, BS) < d_0 : direct\ transmission\ using\ free\ space\ model$
    - $if\ d(CH, BS) \geq d_0 : multi\ hop\ routing\ via\ neighboring\ CHs$

Energy consumption is updated using first-order radio model [18].

- $$\mathrm{E_{res}(i)} = \mathrm{E_{res}(i)} - \mathrm{E_{tx}} - \mathrm{E_{rx}} - \mathrm{E_{agg}} \quad (11)$$
- if $\mathrm{E_{res}(i)} < 0$, the node is marked as dead

Where $E_{agg}$is the aggregation energy, $E_{tx}$is transmission energy, and E_{rx} is receiving energy.

I. Pseudocode

Algorithm 1: AdaHet-Clust

```
for each round r = 1 to MAX_ROUNDS do
    N_alive ← |{i: E_res(i)>0}|; if N_alive=0 then break
    E_total_res ← Σ E_res(i)
    // Phase 1-2: Setup & CH Election (parallel)
    for each node i do
        E_avg ← (E_res(i)+ΣE_res(neighbors))/(1+|neighbors|)
        α ← E_max(i)/max(E_max(neighbors))
        β ← 1 - d(i,BS)/d_max
        γ ← N_alive/N
        T_adapt ← T_base × E_total_res/E_total_init
        P_ch ← P_opt × (E_res(i)/E_avg) × α × β × γ
        if P_ch > T_adapt and random () <P_ch then
            is_CH(i) ← true; broadcast (CH_ANNOUNCEMENT)
    // Phase 3: Collision Resolution
    for each CH i do
        if multiple CHs within 2 hops, then
            keep node with max (E_res/(1+deg)) as CH
    // Phase 4: Cluster Formation
    Non-CHs join nearest CH; CHs create TDMA schedules
    // Phase 5: Data Transmission
    CHs: collect → aggregate → forward to BS
    Non-CHs: transmit to CH in TDMA slot
    Update E_res(i); mark dead if E_res(i)≤0
end for
return FND, HND, LND
```

IV. THEORETICAL ANALYSIS

This section provides a comprehensive theoretical analysis of AdaHet-Clust, including time

complexity, message complexity, space complexity, energy bounds, and scalability properties.

Theorem 1 (Time Complexity): Each round of AdaHet-Clust terminates in *O(1)* time steps.

Proof: All five phases have bounded duration with no iterative refinement.

Theorem 2 (Message Complexity): In each round, the total number of control messages is at most 3N.

Proof: Each node transmits at most one energy broadcast (amortized), one CH announcement, and one join request.

Theorem 3 (Space Complexity): Each node requires *O(deg(i))* storage.

Proof: Nodes store neighbor tables of *size O(deg(i))* and TDMA schedules of the same order

Theorem 4 (Energy Bound per Round): The control energy consumed per round is bounded by:

The control energy consumed per round is bounded by:

$$E_{control} \leq N * (E_{tx}^{ctrl} * (1 + P_{ch}^{max}) + E_{rx}^{ctrl} * d') \quad (12)$$

Where E_{rx}^{ctrl} *and* E_{tx}^{ctrl} are the energies for receiving/transmitting a control message, P_{ch}^{max} less or equal 1 is the maximum CH probability, and d' is the average node degree

Proof: Each node transmits ≤ *1* control message if running for CH (probability ≤ P_{ch}^{max} ≤1) and 1 join request. Control messages are short. Thus $E_{control} \ll E_{data}$.

Theorem 5 (Energy-Proportional CH Election): For any two neighboring node *i* and node *j*, the probability of their election in CH ratio the following condition:

$$\frac{Pr[i \text{ is CH}]}{Pr[j \text{ is CH}]} = \frac{E_{res}(i)}{E_{res}(j)} * \frac{E_{max}(i)}{E_{max}(j)} * \frac{\beta(i)}{\beta(i)} \quad (11)$$

Proof: from (4), $P_{ch}(i)$ for neighboring nodes, $\gamma(r)$ *and* E_{avg_local} are approximately equal.

Substituting $\alpha(i) = \frac{E_{max}(i)}{M}$ and $\alpha(j) = \frac{E_{max}(j)}{M}$ where $M = \max_{k \in N_1(i) \cup \{i\}} E_{max}(k) \approx \max_{k \in N_1(j) \cup \{j\}} E_{max}(k)$ for neighbors, the ratio simplifies to (11).

Theorem 6 (Convergence): AdaHet-Clust terminates each round in *O(1)* time steps with probability 1.

Theorem 7 (Scalability): The performance degradation of AdaHet-Clust as network size increases is bounded by *O(logN*) in terms of HND. Table II: Complexity Comparison with Baseline Algorithms, Table II: illustrate the Complexity Comparison**.**

TABLE II:
COMPLEXITY COMPARISON WITH BASELINE ALGORITHMS

Complexity Metric	Time per round	Messages per round	Per-node space	Computation per node	Scalability
LEACH	$O(1)$	$O(N)$	$O(1)$	*Low*	*Good*
HEED	$O(1)$	$O(N)$	$O(deg)$	*Low*	*Good*
DEEC	$O(1)$	$O(N)$	$O(deg)$	*Low*	*Good*
HMGEAR	$O(1)$	$O(N^2)$	$O(N)$	*Moderate*	*Good*
AdaHet-Clust	$O(1)$	$O(N)$	$O(deg)$	*Low*	*Excellent*

V. SIMULATION RESULTS AND DISCUSSION

A. Simulation Setup

we have performed an extensive simulation study using MATLAB R2013a. The network consists of 100 nodes (default) with three different heterogeneities: 50% low (0.5 J), 30% medium (1.0 J), 20% high-energy (2.0 J). Simulation performs until all nodes deplete energy. We compared our proposed work against three protocols: LEACH (Katiyar et al., 2011), DEEC (Ordouei), and HMGEAR (Moasses et al., 2021).

B. Simulation Results

In Fig. 2: shows network lifetime comparison with error bars (±1 standard deviation). Our proposed work achieves FND of 1651 ± 72 rounds, compared with (1420 ± 68, +16.3%), (1108 ± 61, +49.0%), and (612 ± 45, +169.8%), HMGEAR, DEEC, and LEACH respectively. HND reaches 2654 ± 104 compared with (2310 ± 98, +14.9%), (1897 ± 92, +39.9%), and (1218 ± 78, +118.0%), HMGEAR, DEEC, and LEACH. LND reaches 3412 ± 168 versus (2980 ± 152, +14.5%), (2591 ± 135, +31.7%), and (1834 ± 112, +86.0%), HMGEAR, DEEC and LEACH respectively. Non-overlapping error bars confirm statistical significance.

These improvements are attributed to three key mechanisms. First, the adaptive threshold decreases as network energy depletes, enabling more aggressive CH election when needed. This prevents low-energy

nodes from becoming CHs early while ensuring adequate CH coverage later, an advantage over HMGEAR's static CH selection which cannot adapt to energy depletion patterns. Second, the dynamic heterogeneity factor continuously prioritizes high-capacity nodes based on local energy ratios, unlike DEEC's static weights and HMGEAR's gateway-centric approach which cannot adapt to node-level energy differences. Third, the two-hop deterministic collision resolution ensures optimal CH placement without contention, whereas HMGEAR relies on gateway coordination which can create bottlenecks.

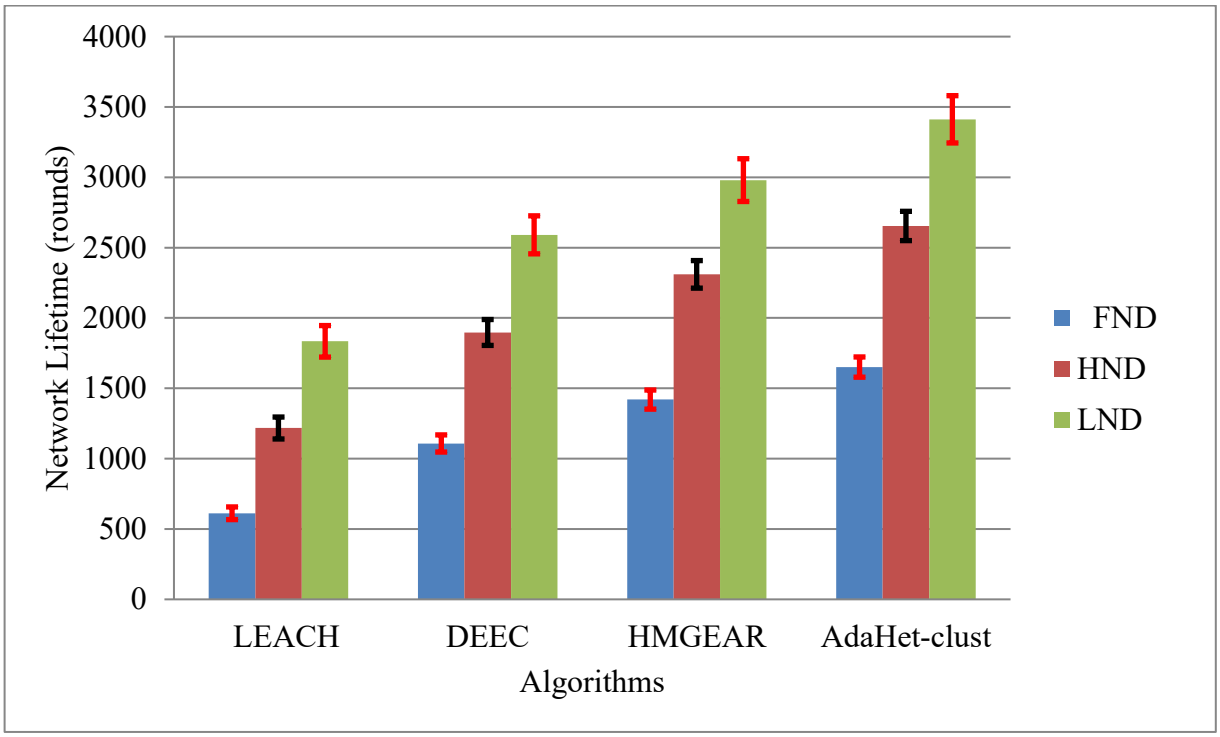


Fig. 2: Network lifetime comparison (FND, HND, LND) with error bars showing plus or minus one standard deviation

In Fig. 3: shows the energy distribution uniformity at the Half Node Death point, measured as the standard deviation of residual energy among alive nodes. Lower standard deviation indicates fairer energy distribution. Shows residual energy std dev at HND. AdaHet-Clust achieves 0.149 J, versus HMGEAR (0.175 J, -14.9%), DEEC (0.221 J, -32.6%), and LEACH (0.452 J, -67.0%).

The wide spread in LEACH confirms that random CH selection without energy awareness leads to highly uneven energy consumption. HMGEAR's gateway-based approach provides better uniformity than DEEC due to its residual energy-aware CH selection, but it cannot match AdaHet-Clust's dynamic adaptation because HMGEAR relies on static CH selection and gateway coordination rather than continuous node-level adaptation based on local energy ratios.

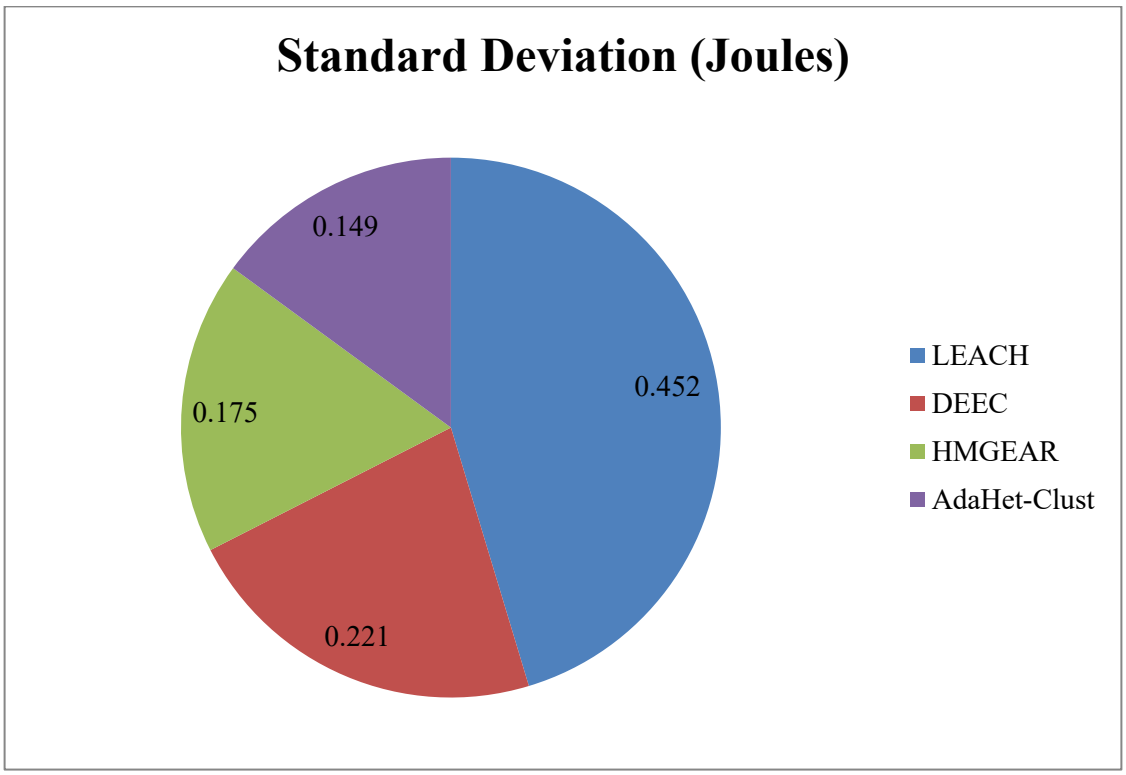


Fig. 3: Energy uniformity comparison at Half Node Death.

The enhanced energy distribution achieved by the suggested algorithm has significant practical implications for network reliability and longevity. By balancing energy consumption more uniformly between sensor nodes, the algorithm mitigates the premature depletion of individual nodes, thereby decreasing the likelihood of coverage holes and network partitioning. The dynamic heterogeneity factor continuously adjusts CH probabilities based on local energy ratios, effectively balancing the load across all nodes regardless of their initial capacity.

Fig. 4: compares the control overhead measured in kilobytes transmitted per round. This metric is critical for energy-constrained IoT networks where communication often dominates energy consumption.

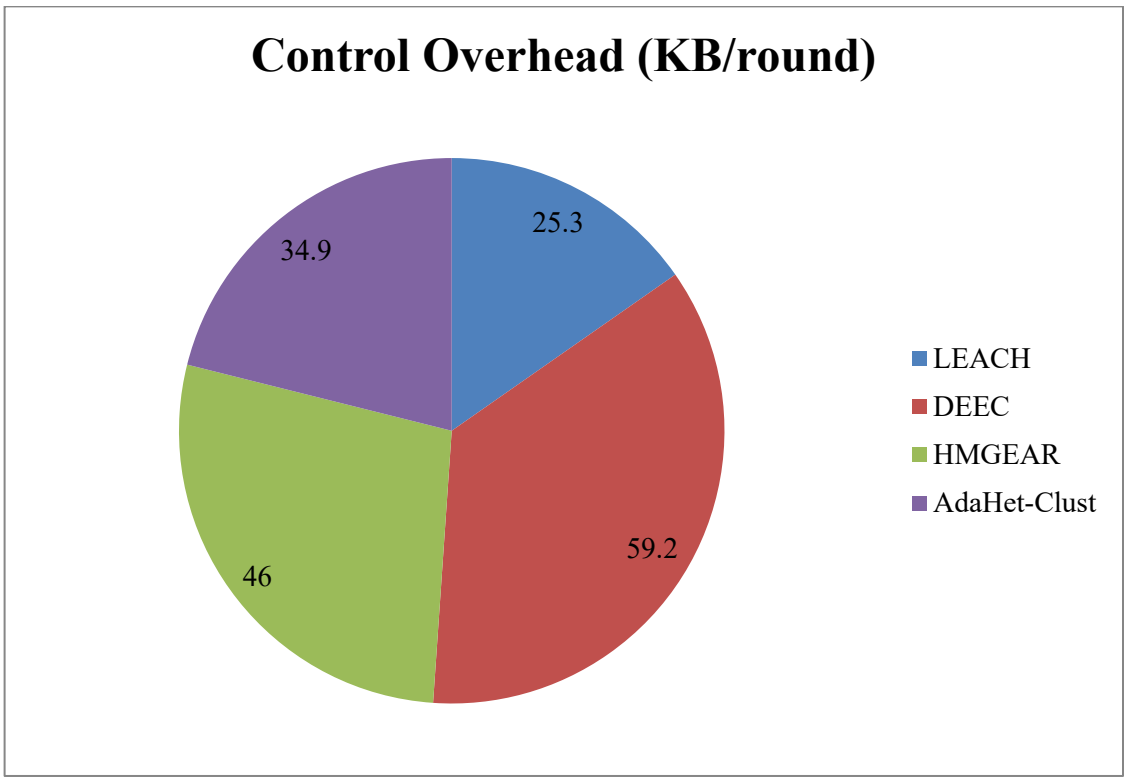


Fig. 4: Control overhead comparison (KB per round)

AdaHet-Clust requires only 34.9 KB per round, which is 24.1% less than HMGEAR (46.0 KB), 41.1% less than DEEC (59.2 KB), and only 38.0% higher than LEACH (25.3 KB). LEACH has the lowest overhead but at the cost of significantly worse network lifetime (169.8% lower FND).

HMGEAR's gateway coordination adds overhead for synchronization between sensor nodes and the gateway, while DEEC's round-based probability decay requires periodic recalculations and global information collection.

The overhead reduction compared to HMGEAR stems from AdaHet-Clust's fully distributed operation without requiring gateway coordination or a dedicated gateway node. The single-round CH election process eliminates the multiple iterations required by DEEC, and no global information collection is needed unlike HMGEAR's gateway-centric approach which requires data forwarding to a central gateway for aggregation.

In Fig. 5: presents the scalability analysis, showing how HND varies as network size increases from 50 to 500 nodes. All algorithms exhibit performance degradation as network size increases due to increased contention, but AdaHet-Clust maintains the largest absolute advantage.

AdaHet-Clust achieves at N=00, 2120 ± 142 rounds compared with (1890 ± 135, +12.2%), (1500 ± 122, +41.3%), and (950 ± 112, +123.2%), for HMGEAR, DEEC, and LEACH respectively. Non-

overlapping error bars at 200 nodes and 500 nodes between AdaHet-Clust and all baselines confirm statistical significance.

The degradation from nodes =50 to 500 is 27.1% for AdaHet-Clust, versus 26.8%, 28.6%, and 29.6%, for HMGEAR, DEEC, and LEACH respectively. While HMGEAR shows better relative degradation due to its gateway assistance, AdaHet-Clust maintains a significantly higher absolute HND across all network sizes without requiring any gateway infrastructure. HMGEAR's gateway becomes a bottleneck as network size increases, limiting its absolute performance.

That shows optimal P_{opt} =0.05 achieves HND = 2654 ± 104. At P_{opt} =0.04 (2580 ± 98) and 0.06 (2610 ± 102), overlapping error bars indicate no significant difference. Performance varies ±2.5% within 0.04-0.06. The performance degradation on either side of the optimum is modest. Within the range of 0.04 to 0.06, HND varies by only ±2.5%. This indicates that AdaHet-Clust is robust to parameter misconfiguration. The optimal value of 0.05 corresponds to having approximately 5% of nodes serve as CHs, which balances adequate coverage against control overhead.

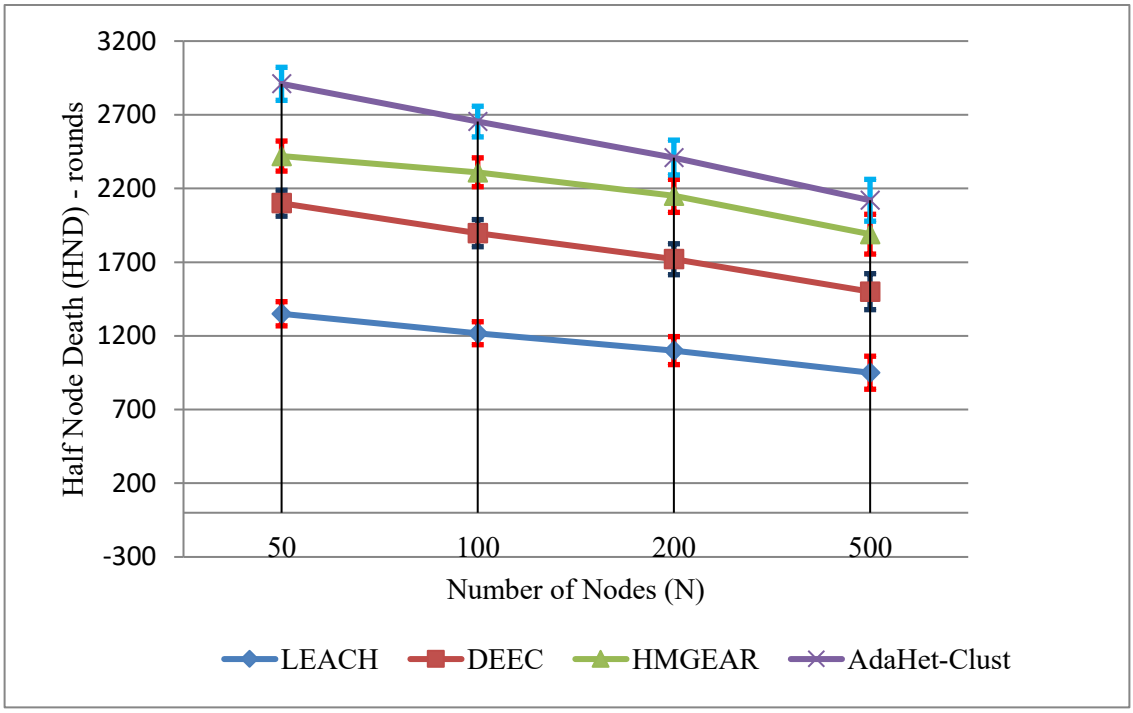


Fig. 5: Scalability analysis: HND versus network size

In Fig. 6: examines the sensitivity of AdaHet-Clust to the parameter Popt, which determines the target percentage of nodes that should serve as CHs.

That shows optimal P_{opt} =0.05 achieves HND = 2654 ± 104. At P_{opt} =0.04 (2580 ± 98) and 0.06 (2610 ± 102), overlapping error bars indicate no significant difference. Performance varies ±2.5% within 0.04-

0.06. The performance degradation on either side of the optimum is modest. Within the range of 0.04 to 0.06, HND varies by only ±2.5%. This indicates that AdaHet-Clust is robust to parameter misconfiguration. The optimal value of 0.05 corresponds to having approximately 5% of nodes serve as CHs, which balances adequate coverage against control overhead.

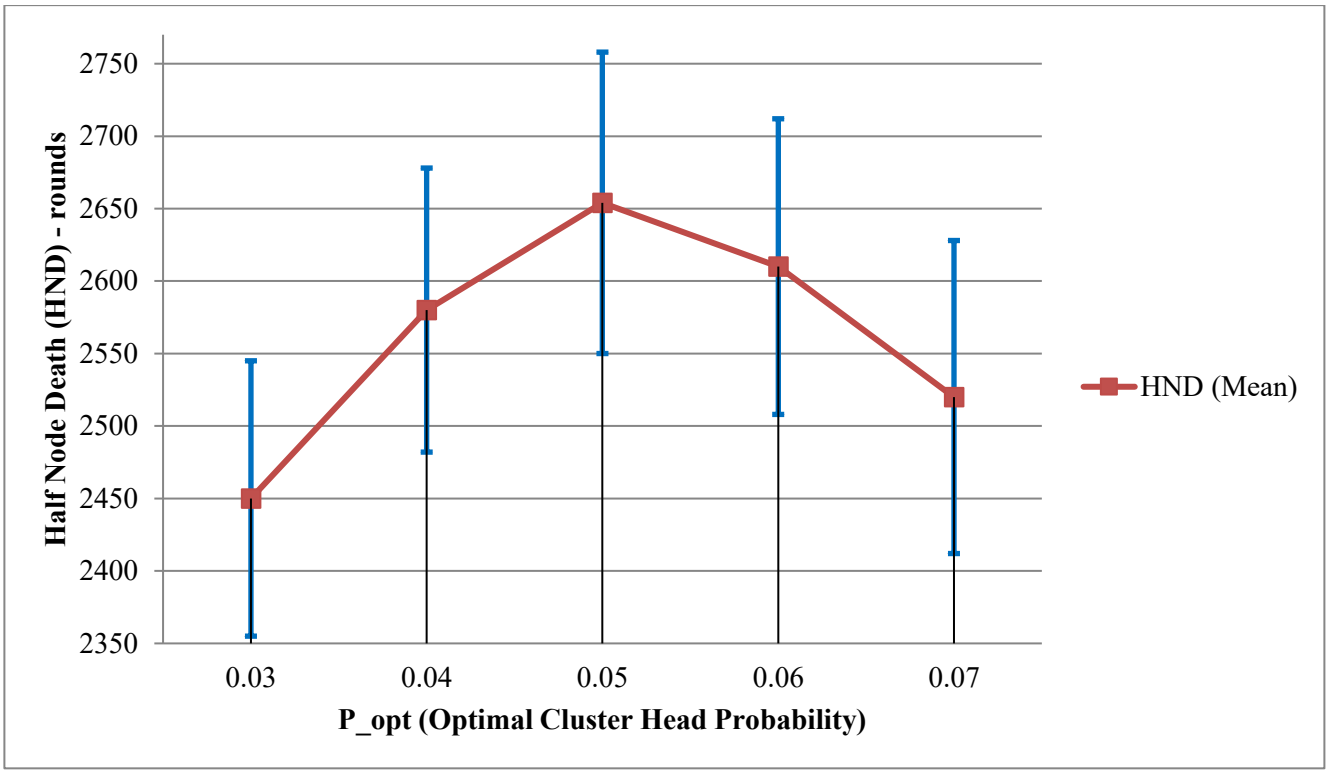


Fig. 6: Sensitivity to base threshold.

Fig. 7: shows the sensitivity to T_{base} (HND in rounds (N=100)), which controls the initial threshold before network energy depletion begins to lower it.

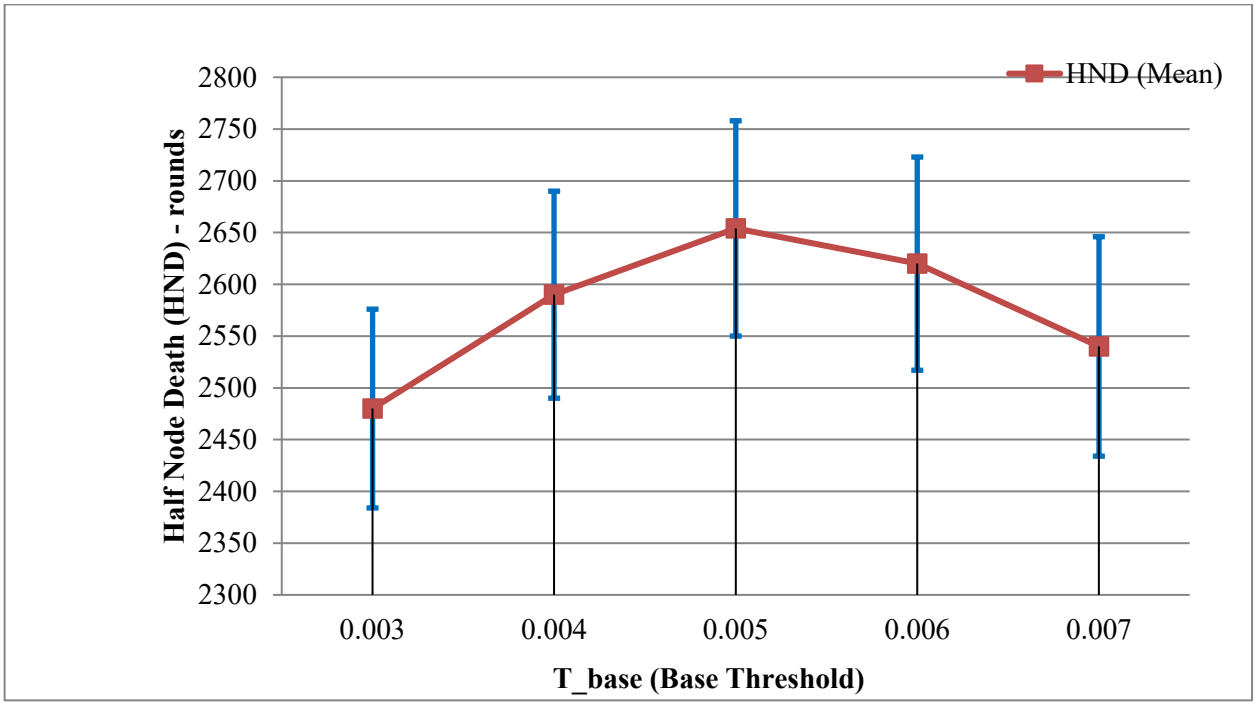


Fig. 7: Sensitivity to heterogeneity ratio under three scenarios.

HND achieved 2480 rounds at T_{base}=0.003. HND increase to (2590 ± 100) rounds at T_{base}=0.004. . The optimal performance is obtained at T_{base}=0.005, achieving a Half Node Death (HND) of (2654 ± 104) rounds, and the T_{base} is increasing to 0.006 results in a slight decrease in HND to (2620±103) rounds, while a further increase to 0.007 decreases HND to (2540±106) rounds, overlapping error bars confirm robustness.

Notably, the optimal value T_{base}=0.005 corresponds to 10% of the optimal cluster head election, p_{opt} =0.05. This proportional relationship is intuitive: the base threshold should be set to a fraction of the target CH probability to provide a conservative baseline that becomes more aggressive as energy depletes.

In Fig. 8: investigates how different distributions of node energy capacities affect algorithm performance. AdaHet-Clust is evaluated by 3 scenarios are: low, medium, and high heterogeneity. The AdaHet-Clust in low heterogeneity scenario achieves 2350 rounds HND, versus HMGEAR, DEEC, and LEACH (2050, +14.6%), (1720, +36.6%), and (1180, +99.2%) respectively. Our proposed schema achieves in the default scenario 2654 rounds compared with HMGEAR, DEEC, and LEACH (2310, +14.9%), (1897, +39.9%), and (1218, +118.0%) respectively. In finally our proposed schema in the high heterogeneity scenario achieves 3100 rounds compared with HMGEAR, DEEC, and LEACH (2710, +14.4%), (2200, +40.9%), and (1350, +129.6%) respectively.

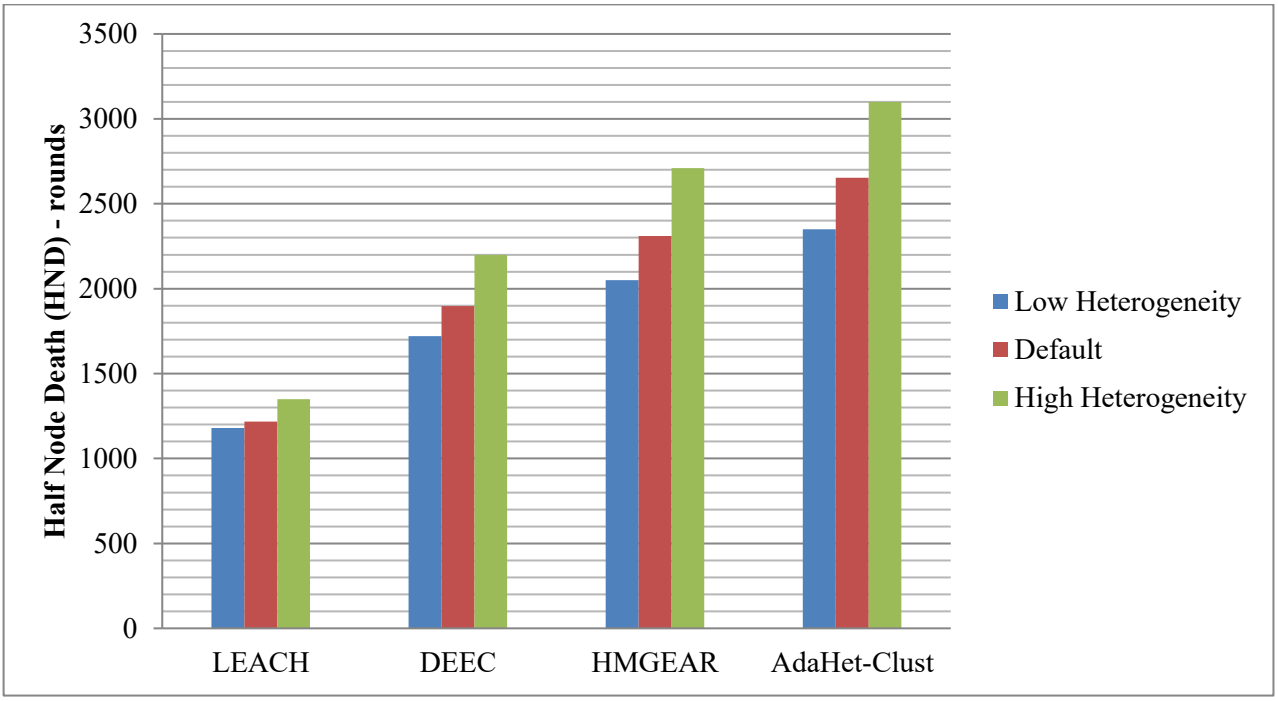


Fig. 8: Sensitivity to heterogeneity ratio under three scenarios.

The most important observation is that AdaHet-Clust's advantage over HMGEAR remains strong across all heterogeneity levels (14.4-14.9%), while its advantage over DEEC increases with heterogeneity (36.6% to 40.9%). This confirms that the dynamic heterogeneity factor $\alpha(i)$ effectively leverages high-capacity nodes, while HMGEAR's gateway-based approach and DEEC's static weights cannot adapt to distribution changes. Unlike HMGEAR which relies on gateway coordination, AdaHet-Clust's dynamic adaptation works locally without additional infrastructure.

C. Discussion

The results confirm that AdaHet-Clust effectively addresses the limitations of existing clustering algorithms. The 16.3% FND improvement over HMGEAR and 49.0% over DEEC validates that the adaptive threshold mechanism prevents premature low-energy node death while the dynamic heterogeneity factor ensures high-capacity nodes carry the energy burden.

AdaHet-Clust achieves 169.8% higher FND with only 38% higher control overhead a highly favorable trade-off. AdaHet-Clust provides 39.9% higher HND while reducing control overhead by 41.1%, proving dynamic adaptation superior to static weighting. While HMGEAR uses a dedicated gateway node, AdaHet-Clust achieves 16.3% higher FND, 14.9% higher HND, and 24.1% lower control overhead without any gateway infrastructure, eliminating single point of failure and hardware costs.

Performance varies by only ±2.5% within optimal parameter ranges, confirming practical deploy

ability without fine-tuning. AdaHet-Clust maintains 12-16% higher HND than HMGEAR from 50 to 500 nodes without gateway bottlenecks. AdaHet-Clust's advantage over HMGEAR remains strong across all heterogeneity levels (14.4-14.9%), confirming that dynamic α(i) effectively leverages high-capacity nodes where gateway-based approaches cannot.

In summary, AdaHet-Clust outperforms all baselines across all metrics while requiring no gateway, and no centralized coordination making it uniquely suitable for resource-constrained heterogeneous IoT deployments.

VI. CONCLUSION

This paper presented AdaHet-Clust, a lightweight, fully distributed, energy-aware clustering algorithm for heterogeneous IoT sensor networks using adaptive thresholding. The key innovations are fourfold. First, a dynamic heterogeneity factor that prioritizes higher-capacity nodes without static weights, enabling effective load balancing across nodes with different energy capacities. Second, a real-time adaptive threshold that responds to network energy depletion, becoming more aggressive as energy declines to maintain cluster head coverage. Third, a deterministic two-hop collision resolution mechanism using score-based ranking, ensuring optimal cluster head placement without contention.

Fourth, comprehensive theoretical analysis is presented through seven formal theorems that establish the algorithm's computational efficiency and performance characteristics, including time complexity O(1), message complexity O(N), space complexity O(deg(i)), scalability, convergence, energy bounds, and energy proportionality.

The simulation results show that AdaHet-Clust significantly outperforms the benchmark protocols in terms of energy balancing, network lifetime, and lower control overhead. Robustness to parameter variations (±2.5%) and maintains consistent performance across network sizes ranging from 50 to 500 nodes, make it suitable for practical for IoT deployments. Future research direction includes extending this work to support mobility, incorporating multi-hop intra-cluster routing for large area deployments, energy harvesting integration into the network health factor, developing lightweight secure clustering with authentication.

REFERENCES

Ab Rahman, A. and Hamim, S. I., 2025. Improved K-means clustering and adaptive distance threshold for energy reduction in WSN-IoTs. *Array,* 27**,** 100436.

Al-Fuqaha, A., Guizani, M., Mohammadi, M., Aledhari, M. and Ayyash, M. 2015. Internet of things: A survey on enabling technologies, protocols, and applications. *IEEE communications surveys & tutorials,* 17**,** pp.2347-2376.

Al-Hussain, E. A. and Al-Suhail, G. A. 2025. Energy-Aware Clustering Using Intelligent Scheme for Heterogeneous Wireless Sensor Networks. *Karbala International Journal of Modern Science,* 11**,** 3.

Alomari, M. F., Mahmoud, M. A. and Ramli, R. 2022. A systematic review on the energy efficiency of dynamic clustering in a heterogeneous environment of wireless sensor networks (wsns). *Electronics,* 11**,** 2837.

Alsharari, T., Alresheedi, S., Fatani, A. and Maolood, I. 2020. Significant role of internet of things (IoT) for designing smart home automation and privacy issues. *International Journal of Engineering and Technology,* 9**,** pp.515-519.

Askar, N. A., Habbal, A., Alden, F. Z., Wei, X., Alaidaros, H., Guo, J. and Yu, H. 2023. Forwarding strategies for named data networking based IoT: Requirements, taxonomy, and open research challenges. *IEEE access,* 11**,** pp.78363-78383.

Askar, S. K. 2023. Deep forest based internet of medical things system for diagnosis of heart disease. *aro-the scientific journal of koya university,* 11**,** pp.88-98.

Atzori, L., Iera, A. and Morabito, G. 2010. The internet of things: A survey. *Computer networks,* 54**,** pp.2787-2805.

Baranidharan, B. and Santhi, B. 2015. GAECH: genetic algorithm based energy efficient clustering hierarchy in wireless sensor networks. *Journal of Sensors,* 2015**,** 715740.

Darabkh, K. A., Al-Maaitah, N. J., Jafar, I. F. and Khalifeh, A. F. 2018. EA-CRP: a novel energy-aware clustering and routing protocol in wireless sensor networks. *Computers & Electrical Engineering,* 72**,** pp.702-718.

Dwivedi, A. K. and Sharma, A. K. 2020. NEEF: A novel energy efficient fuzzy logic based clustering protocol for wireless sensor network. *Scalable Computing: Practice and Experience,* 21**,** pp.555-568.

Ekler, P., Levendovszky, J. and Pasztor, D. 2022. Energy aware IoT routing algorithms in smart city environment. *IEEE Access,* 10**,** pp.87733-87744.

Farooq, N. and Abdul, M. 2018. A survey on energy efficient clustering protocols in wireless sensor networks. *International Journal of Computer Applications,* 180**,** pp.9-12.

Goyal, A., Mudgal, S. and Kumar, S. A review on energy-efficient mechanisms for cluster-head selection in WSNs for IoT application. IOP Conference Series: Materials Science and Engineering, 2021. IOP Publishing, 012010.

Haris, M. and Nam, H. 2025. Enhancing energy efficiency in iot-wsns through optimized pso cluster head selection. *IEEE Access*.

Heinzelman, W. R., Chandrakasan, A. and Balakrishnan, H. Energy-efficient communication protocol for wireless microsensor networks. Proceedings of the 33rd annual Hawaii international conference on system sciences, 2000. IEEE, 10 pp. vol. 2.

Ibriq, J. 2004. Cluster-based routing protocol in wireless sensor networks: Issues and challenges. *Proc. SPECTS, 2004***,** pp.759-766.

Javaid, N., Qureshi, T., Khan, A., Iqbal, A., Akhtar, E. and Ishfaq, M. 2013. EDDEEC: Enhanced developed distributed energy-efficient clustering for heterogeneous wireless sensor networks. *Procedia computer science,* 19**,** pp.914-919.

Jibreel, F., Tuyishimire, E. and Daabo, M. 2022. An enhanced heterogeneous gateway-based energy-aware multi-hop routing protocol for wireless sensor networks. Information (2022). s Note: MDPI stays neutral with regard to jurisdictional claims in published ….

Katiyar, V., Chand, N. and Soni, S. 2011. A survey on clustering algorithms for heterogeneous wireless sensor networks. *situations,* 19**,** 20.

Khorsheed, O. K. 2025. Mitigating Dead Node Impact on Coverage and Connectivity in Wireless Sensor Networks Using a Hybrid Approach. *ARO-THE SCIENTIFIC JOURNAL OF KOYA UNIVERSITY,* 13**,** pp.131-143.

Lilo, M. A., Yasari, A. K., Hamdi, M. M. and Abbas, A. D. 2024. Transmission Power Reduction Based on an Enhanced Particle Swarm Optimization Algorithm in Wireless Sensor Network for Internet of Things. *Aro-The Scientific Journal Of Koya University,* 12**,** pp.61-69.

Lonkar, B., Charde, P., Kuthe, A., Dehankar, A., Dandekar, P., Puri, N. and Banabakode, D. 2026. A review of hybrid clustering and routing protocols for the hotspot problem in wireless sensor networks. *Peer-to-Peer Networking and Applications,* 19**,** 8.

Moasses, H., Ghaderzadeh, A. and Khamforoosh, K. HetEng: An Improved Distributed Energy Efficient Clustering Scheme for Heterogeneous IoT Networks. 2021 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT), 2021. IEEE, pp.79-84.

Mohammed, Q. A. H., Askar, N. A., Habbal, A., Sajat, M. S. and Yusupov, Z. 2022. Architecture Protocols and Applications of the Internet of Medical Things (IoMT). *J. Commun,* 17.

Nehra, V. and Sharma, A. K. 2013. PEGASIS-E: power efficient gathering in sensor information system extended. *Global Journal of Computer Science and Technology,* 13**,** pp.15-18.

Ordouei, M. Energy-Efficient Routing in IoT Sensor Networks Using a Squirrel Optimization Algorithm-Based Clustering Framework.

Ossama, Y. 2004. Heed: a hybrid, energy-efficient, distributed clustering approach for ad-hoc sensor networks. *IEEE Trans. on Mobile Computing,* 3**,** 660-669.

Qing, L., Zhu, Q. and Wang, M. 2006. Design of a distributed energy-efficient clustering algorithm for heterogeneous wireless sensor networks. *Computer communications,* 29**,** 2230-2237.

Reina, D. G., Toral, S. L., Barrero, F., Bessis, N. and Asimakopoulou, E. 2013. The role of ad hoc networks in the internet of things: A case scenario for smart environments. *Internet of things and inter-cooperative computational technologies for collective intelligence.* Springer.

Sharma, T., Goyal, G., Tiwari, B., Tomar, G. and Rathee, N. 2026. Energy-aware Clustering and Lifetime Extension in WSNs Using Evolutionary Learning Algorithms. *International Journal of Computational Intelligence Systems,* 19**,** 1.

Sun, H., Li, Y., Lee, H. and Chen, H. 2026. An efficient clustering protocol driven by hybrid intelligence for improving energy efficiency in industrial internet of things. *Journal of King Saud University Computer and Information Sciences*.

Ti, Y.-W., Cheng, R.-H., Wei, S. and Yu, C.-M. 2026. HEOCP: Hybrid Energy-Optimized Clustering Protocol for WSNs Using Analytical Modeling and Deep Learning Integration. *Sensors,* 26**,** 1188.

Zhu, C., Leung, V. C., Shu, L. and Ngai, E. C.-H. 2015. Green internet of things for smart world. *IEEE access,* 3**,** pp.2151-2162.