



"Improved Algorithmic Approach to Cluster Formation in Wireless Sensor Networks" (IAACF)

*Abdulkadir Ahmad Yako¹, Dahiru Sani Shuaibu²

¹ Kano State Polytechnic, Kano, Nigeria.

² Bayero University Kano, Kano, Nigeria.

DOI: 10.5281/zenodo.18358564

Submission Date: 10 Dec. 2025 | Published Date: 24 Jan. 2026

*Corresponding author: [Abdulkadir Ahmad Yako](#)

Kano State Polytechnic, Kano, Nigeria.

Abstract

Wireless Sensor Networks (WSNs) consist of spatially distributed sensor nodes that cooperatively monitor physical or environmental conditions such as temperature, pressure, and motion. One of the major challenges in WSNs is limited energy resources, which directly affect the longevity and stability of the network. Existing clustering algorithms often use single criteria such as minimum distance or strongest received signal strength (RSS), which may result in routing data through longer aggregate distances and lead to excessive energy dissipation. This paper presents an Improved Algorithmic Approach to Cluster Formation (IAACF) in WSNs that employs an aggregate distance criterion to optimize cluster formation and minimize overall energy consumption. The proposed model modifies the cluster formation stage of the hetDEEC-3 protocol by integrating aggregate distance computation between the cluster members, cluster heads, and base station. Simulation results using MATLAB R2014a demonstrate that the IAACF scheme significantly reduces energy consumption per round, thereby increasing the network lifetime by approximately 25.66% compared to hetDEEC-3. The proposed approach achieves balanced energy usage across nodes, improved data delivery to the sink, and enhanced stability, making it an effective solution for energy-constrained WSN applications.

Keywords: Wireless Sensor Networks (WSN), Cluster Formation, Energy Efficiency, Aggregate Distance, Heterogeneous Networks, IAACF, hetDEEC-3, MATLAB Simulation.

Chapter One

1. Introduction

A sensor node is a small/tiny device comprising three basic components: sensing, processing, and wireless communication components (Muhammad et al., 2016), and a typical sensor node configuration is shown in Fig. 1.1.

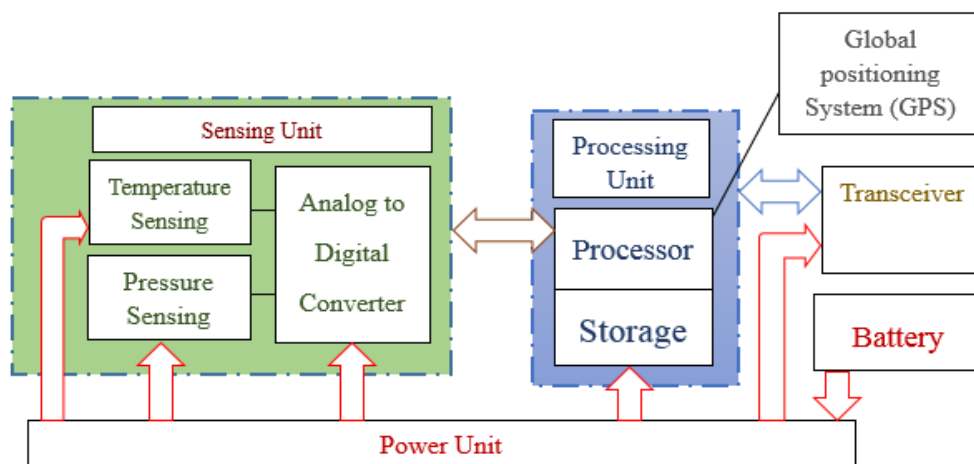


Fig. 1.1 Basic components of a sensor node (Muhammad et al., 2016)

Wireless Sensor Network (WSN) is a popular medium of low-cost infrastructure communication and is slowly emerging as a wireless technology among various classes of communication networks, such as Cellular Networks, Adhoc Networks, and Mesh Networks (Akila & Baby, 2016). Typically, an ad hoc network consists of a number of nodes that communicate via wireless connection without any aid from centralized administrative control (Saleh Ali & Sumari, 2010). The Wireless Sensor can act as a sensor node, sensing physical/environmental phenomenon and sending the sensed data to a remote server or base station (Central Gateway) for further action. A WSN consists of spatially distributed independent sensor nodes that collectively monitor physical or surrounding conditions in an organized manner, and the nodes interact wirelessly and cooperatively in an ad-hoc fashion after being deployed in an environment to monitor certain physical conditions. Depending on the application realm, WSN may contain hundreds or even thousands of nodes, and the sensing nodes communicate their sensed data to the destination node across an intermediate node, with this destination node interconnected to a central gateway, also referred to as base station or sink node.

1.1 Motivation and Significance of the Study

This study is motivated by the necessity of having optimized WSNs that will be efficient in power/energy consumption and more durable in terms of life-span. The study is conceived to enhance the preservation of energy and balance the power consumption of the sensor nodes, which will lead to improvement in the lifespan of the WSN. WSNs can be used in different surroundings for monitoring tasks such as disaster relief, rescue, search, and target tracking, and obtaining an efficient power consuming WSN with guarantee of battery life makes it feasible and beneficial even in highly constrained environments. Sometimes these sensor nodes are deployed in hazardous areas like volcanic, battle field, water flood, and underwater for monitoring natural phenomenon or human activities, and in these positions, it is very difficult if not impossible to recharge or replace the sensor nodes. Hence, a robust and efficient power saving algorithm is needed so that energy consumption is minimized to enhance and maximize the lifespan of the network. Therefore, the need to have an effective energy saving algorithm that will not only elongate the lifespan of the network but also reduce latency by taking effective route in data transmission serves as motivation for this study, while the significance of the study is to show that the cluster formation criterion of a WSN can be improved to save precious energy of the nodes which will in turn translate into increasing the lifetime of the network.

1.2 Wireless Sensor Networks

The central gateway offers an additional connection to the wired/wireless world where the data can be harvested, processed, and analyzed, and Fig. 1.2 shows how the WSN is connected to a remote end user via internet.

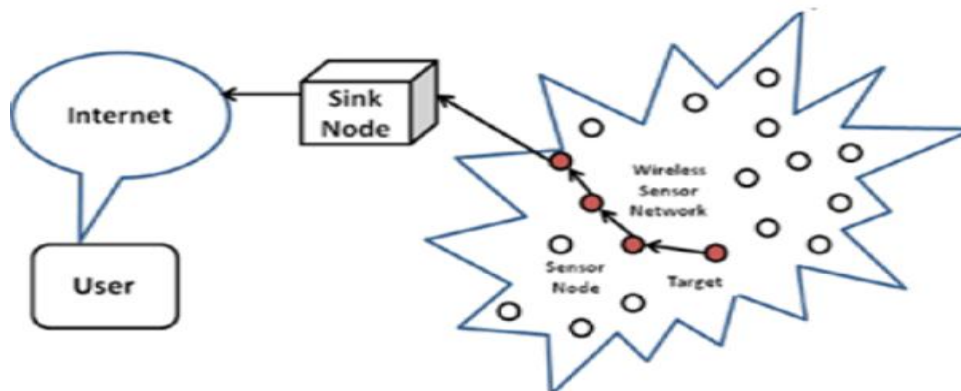


Fig. 1.2 Connection of the WSN to a remote end user (Akila and Baby, 2016)

Therefore, a WSN could be regarded as a bunch of tiny, lightweight, portable sensor nodes stationed to supervise assets, battle fields, transport systems, and surroundings (Akila & Baby, 2016). To attain higher degree of precisions, these sensor nodes are haphazardly stationed in the concerned region or very near to the concerned region (Thiriveni & Ramakrishnan, 2016). One of the constituents of a sensor node is the power supply which sources the energy required by the sensor to affect the scheduled tasks. These sensors may have standardized characteristics defining the homogeneity known as homogeneous WSNs, and if these sensors differ in price, storage, communications method, computational disparities, sensing styles, and energy capabilities, then they are said to be heterogeneous WSNs. WSNs were foreseen to become the texture of our community and surroundings; nevertheless, they are still incapable of overcoming numerous operational difficulties such as restricted energy of the sensor nodes which suffocates their wide-spread emplacement (Nadeem et al., 2015). There are many types of WSN topologies being used by many researchers, and according to Divya et al. (2013) these include Tree Topology, Bus Topology, Ring Topology, Star Topology, and Circular Topology. However, the main challenges of these WSNs include the limited battery capacity carried by nodes due to their size, which is impossible to replace after deployment, and the nodes are restricted by having 8-bit and 16-bit microcontroller as processor with low output transmission power resulting in short communication range.

1.3 Problem Statement

In most of the existing works reviewed in this study, the cluster members use either minimum distance or strongest received signal strength as their cluster formation criterion, which may lead to routing the sensed data through a longer aggregate distance to the sink as shown in Fig. 1.3.

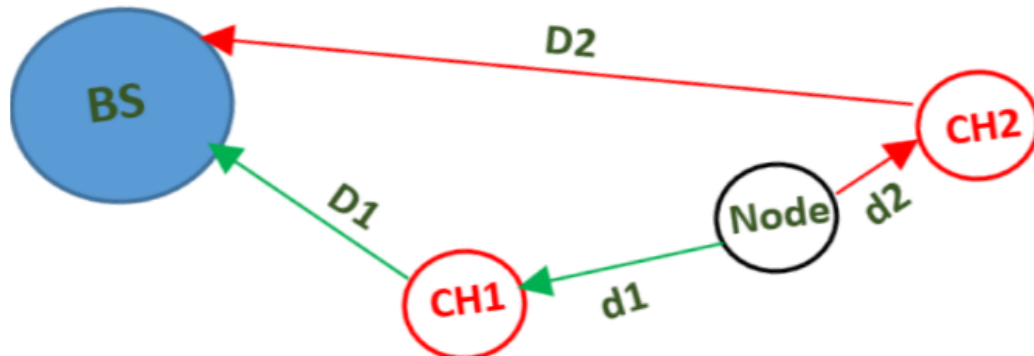


Fig. 1.3: Cluster formations in WSN

Fig. 1.3 shows that if the criterion of minimum distance or strongest received signal strength is used, the node in question will automatically choose CH2 to send its sensed data to the BS. This means that the sensed data will have to go through longer aggregate distance to the BS, resulting in dissipating more of the much-needed energy that could be used to enhance the lifespan of the network. It can also be seen that the route through CH1 is the shortest route, having the smallest aggregate distance as compared to the route through CH2, and it is well known that in WSN the distance travelled by data is directly proportional to the energy consumed by the network, meaning the shorter the route the less the energy consumed in transmitting the data to various destinations. Samayveer et al. (2016) use the strongest received signal strength in the cluster formation process of their work, which may mean routing the sensed data through a longer aggregate distance to the BS as seen in Fig. 1.3. Based on the foregoing observations, therefore, this study focused on modifying the cluster formation criterion of hetDEEC-3 by Samayveer et al. (2016), with intent to develop an algorithm of low power consumption, taking into consideration the concept of heterogeneity in WSNs.

1.4 Aim and Objectives

The aim of this study is to develop an IAACF for WSN. The following objectives are set to be achieved by the research:

1. To modify the cluster formation algorithm of the hetDEEC-3 protocol of Samayveer et al. (2016).
2. To simulate the modified cluster formation algorithm using Matlab R2014a software.
3. To compare the performances of the modified algorithm with the hetDEEC-3

Chapter Two

2.0 Literature Review

In this section, existing literature was reviewed with emphasis on critically examining the cluster formation criterion used by researchers in their respective works.

2.1 RSS-Based Cluster Head Selection

The researchers that used RSS or received signal strength indicator (RSSI) as a cluster head selection criterion are summarized in Table 2.1.

Table 2.1: Summarized researches where RSS was used as cluster head selection criterion

| S/N | Authors/Year | Title | Cluster Formation Criterion | Criticism |
|-----|--|---|-------------------------------|---|
| 1 | Lonkar, B., Kuthe, A., Charde, P., Dehankar, A., & Kolte, R. /2024 | Hybrid Energy-Saving Cluster Head Selection for WSNs | Minimum distance + RSS hybrid | Proposes a hybrid rule that balances connectivity (RSS) and local energy efficiency (distance), reducing packet drops and improving stability period. |
| 2 | Jha, V., & Sharma, R. /2022 | Energy-Efficient Clustering in Heterogeneous WSNs Using RSS | Strongest RSS | Demonstrates that RSS-based clustering improves connectivity and reduces packet loss in heterogeneous WSNs. |
| 3 | Hein Zelman et al. /2000 | Energy Efficient Communication Protocol for | Received Signal | Better stability, energy consumption rate but silent on packets at BS. |

| | | Wireless Micro Sensor Networks | Strength | |
|----|--------------------------------------|---|--------------------------|--|
| 4 | Kumar, Aseri, & Petal /2011 | A Novel Multihop Energy Efficient Heterogeneous Clustered Scheme for Wireless Sensor Networks | Received Signal Strength | Increases life-span, stability, and greater throughput. |
| 5 | Meenakshi & Sushil /2012 | An Energy Efficient Level Based Clustering Routing Protocol for Wireless Sensor Networks | Received Signal Strength | Better energy, clusters per rounds but silent on packets at BS per rounds. |
| 6 | Asha & Vineeta /2013 | An Election of Vice Cluster Selection Approach to Improve V LEACH Protocol in Wireless Network | Received Signal Strength | Longer lifetime but poor stability. |
| 7 | Monika & Dipak /2014 | An Energy Saving Algorithm to Prolong the Lifetime of Wireless Sensor Network | Received Signal Strength | Compares its result only on network lifetime, neglecting other metrics. |
| 8 | Samayveer Singh /2016 | Energy Efficient Multilevel Network Model for Heterogeneous WSNs | Received Signal Strength | Improves network energy consumption and packets at BS. |
| 9 | Samayveer et al. /2016 | Energy Efficient Heterogeneous DEEC Protocol for Enhancing Lifetime in WSNs | Received Signal Strength | Improves network energy dissipation, aggregate delay, and packets at BS. |
| 10 | Vinith Chauhan & Surender Soni /2019 | Load Balanced Energy Efficient Cluster-chain Based Hybrid Protocol for Wireless Sensor Networks | Received Signal Strength | Improves network energy consumption and packets at BS. |

From the reviewed works, it is evident that RSS or RSSI has been widely adopted as a cluster head selection criterion in WSNs. While several studies demonstrate improvements in connectivity, energy efficiency, and throughput, many of them remain silent on critical performance metrics such as packet delivery at the base station or stability across rounds. Some approaches, such as hybrid methods combining RSS with minimum distance, attempt to balance connectivity with energy efficiency, yet others focus narrowly on network lifetime while neglecting broader evaluation parameters. Overall, the reliance on RSS as a single criterion often leads to limitations in stability, aggregate delay, and comprehensive performance assessment.

2.2 Minimum Distance–Based Cluster Head Selection

While those that used minimum distance as a cluster head selection criterion are also summarized in Table 2.2.

Table 2.2: Summarized researches where minimum distance was used as cluster head selection criterion.

| S/N | Authors/Year | Title | Cluster Formation Criterion | Criticism |
|-----|---|---|---------------------------------|--|
| 1 | Sushil Lekhi & Satvir Singh /2021 | LEACH-Based Energy Efficient Clustering Protocol for WSNs | Minimum distance | Provides a baseline LEACH variant, comparing distance-based clustering local node lifetime to random clustering. |
| 2 | Marcin Lewandowski & Bartłomiej Placzek /2025 | Cluster Head Selection Algorithm for Extending Last Node Lifetime | Distance priority | Improving global longevity compared to earlier distance-only methods. |
| 3 | Zhanyang et al. /2013 | An Energy-Efficient Clustering Routing Algorithm for Heterogeneous Wireless Sensor Networks | Minimum distance of CMs from CH | Longer lifespan, lower energy consumption rate. |
| 4 | Divya et al. /2013 | Increase the Alive Nodes Based on the Cluster Head Selection Algorithm for Heterogeneous | Minimum distance of CMs from CH | Improved lifespan and poor stability in the network. |

| | | Wireless Sensor Networks | | |
|----|------------------------------------|--|---------------------------------|--|
| 5 | Sandeep & Mamta /2014 | Efficient Cluster Head Selection Scheme for Wireless Sensor Network Using Deterministic Protocol | Minimum distance of CMs from CH | Longer stability and shorter lifespan compared to one of algorithms. |
| 6 | Arezoo Abasi & Hedieh Sajedi /2016 | Fuzzy-Clustering Based Data Gathering in Wireless Sensor Network | Minimum distance of CMs from CH | Less number of dead nodes but longer lifetime and increased energy remained. |
| 7 | Tapaswini et al. /2017 | TEEN-V: A Solution for Intra-Cluster Cooperative Communication in Wireless Sensor Network | Minimum distance of CMs from CH | No comparison made to any other network. |
| 8 | Kalu, J. Chinedu /2017 | Development of an Energy Efficient Algorithm to Prolong the Lifetime of Wireless Sensor Networks | Minimum distance of CMs from CH | Improved lifetime, throughput, energy consumption and poor stability period. |
| 9 | Adamu, M. L. /2017 | Development of an Improved Hybrid Low Energy Adaptive Clustering Hierarchy (LEACH) for Wireless Sensor Network | Minimum distance of CMs from CH | Better lifetime, increased packets at BS and poor total energy in the network. |
| 10 | Srividhya V. & Shankar T. /2019 | Energy Reckoning Distance Based Clustering for Spectrum Aware Cognitive Radio Wireless Sensor Networks | Minimum distance of CMs from CH | Longer stability period, better bits transmitted to BS and less channel usage. |

From the reviewed works, it is clear that minimum distance has been widely applied as a cluster head selection criterion in WSNs. Several studies demonstrate that distance-based clustering improves lifespan, throughput, and energy consumption, while others highlight longer stability periods and better packet transmission to the base station. However, many of these approaches also reveal limitations, including poor stability, neglect of comparative evaluation, or inadequate consideration of total energy in the network. Some methods achieve improved local efficiency but fail to address global longevity comprehensively. Overall, while minimum distance provides certain advantages in prolonging network lifetime and reducing energy consumption, its narrow focus often results in trade-offs that limit stability and broader performance metrics.

2.3 Identified Gaps

The review of both RSS-based and minimum distance-based approaches shows that although these criteria improve certain aspects of clustering in WSNs, they often neglect critical parameters such as aggregate distance, latency, and packet delivery at the base station. The reliance on single-criterion approaches results in trade-offs that limit overall network stability and energy efficiency. These gaps directly align with the problem identified in Chapter One, Section 1.2, where the use of minimum distance or strongest received signal strength may route sensed data through longer aggregate distances to the sink, leading to higher energy dissipation and reduced network lifespan. Therefore, the need to modify the cluster formation criterion, as proposed in this study, is justified by the limitations consistently highlighted across existing works.

2.4 Mathematical Model Adopted from the Existing Works

The algorithm is implemented in a two-level heterogeneous environment, where two features of heterogeneities were considered. These are the 10% of the total number of nodes in the network called advanced nodes (m) and the additional energy factor (α) that the advanced nodes have over the remaining normal nodes. Advanced nodes have to become CHs more often than the normal nodes, which is equivalent to a fairness constraint on energy consumption. Initially, each node can become a CH with a probability P_{opt} , that is, the algorithm guarantees that every one of them will become a CH exactly once in every round. Here, P_{opt} is a predetermined percentage of CHs ($P_{opt}=0.05$, meaning 5% of total nodes are initially selected as CHs). Each normal node will have energy E_o , and advanced nodes will have energy $E_o(1+\alpha)$.

The algorithm was divided into rounds, and in these rounds, CHs are elected by using the threshold and the energy priority values of the nodes. The threshold $T(s)$ and the priority P_{adv} values are calculated as in Lawal (2017) by the formulae in (2.3) and (2.4):

$$T(s) = \begin{cases} \frac{P_{adv}}{1 - P_{adv} \times \left(r \bmod \left(\frac{1}{P_{adv}} \right) \right)} & \text{if } n \in G \\ 0 & \text{otherwise} \end{cases} \quad (2.3)$$

$$P_{adv} = \frac{P_{opt}(1 + \alpha)}{1 + \alpha m} \quad (2.4)$$

where r is the current round number, n is the total number of nodes in the network, m is the 10% of the total number of nodes in the network called advanced nodes, α is the extra energy factor for the advanced nodes, and G is the number of sensor nodes that could be selected as CHs at current round r . Only those sensor nodes which have not been selected as CHs in the current epoch are considered. Each eligible node picks a number randomly between 0 and 1 at the beginning of each round. If the number chosen is less than the threshold $T(s)$ and the sensor node S_i belongs to set G , then it becomes a CH; otherwise, it remains a cluster member. The motivation behind the protocol is that using important and efficient parameters for cluster formation would lead to an energy-effective protocol. Equations (2.3) and (2.4) are the CHs selection mathematical model adopted for the realization of the implemented work.

Let E_o be the starting energy level of all the normal nodes, and m the number of the super nodes, which have α times more energy than each of the normal nodes. Accordingly, there are mN super nodes equipped with a startup energy of $E_o(1 + \alpha)$, and $(1-m)N$ normal nodes equipped with startup energy of E_o . Therefore, the total starting energy of all the nodes in the two-level heterogeneous networks is given as the sum of the startup energy of the super nodes E_{osn} and that of the normal nodes E_{onn} , as shown in (2.5), (2.6), and (2.7):

$$E_{osn} = E_o(1 + \alpha) \quad (2.5)$$

$$E_{onn} = N(1 - m)E_o \quad (2.6)$$

where E_{osn} is the initial energy of the super nodes and E_{onn} is the initial energy of the normal nodes.

Hence, the total initial energy of the network E_{total} is given as

$$E_{total} = N(1 - m)E_o + NmE_o(1 + \alpha) = NE_o(1 + \alpha m) \quad (2.7)$$

Equations (2.5), (2.6), and (2.7) are adopted from the works of Sudeep et al. (2015); Thiriveni and Ramakrishnan (2016); Tuba and Meenakshi (2016); and Samayveer Singh, Aruna Malik, and Rajeev Kumar (2016).

The radio energy model that describes an l-bit message transmitted over a distance d is shown in Fig. 2.9 and is adopted in the implemented work as in Monica and Dipak (2014); Guangjie Han, Xu Jiang, Aihua Qian, Joel J. P. C. Rodrigues, and Long Cheng (2014);

Sudeep et al. (2015); Thiriveni and Ramakrishnan (2016); Tuba and Meenakshi (2016); and Samayveer et al. (2016). The radio model is adopted in order to analyze and compare the developed model with Samayveer et al. (2016). Radio model's energy dissipation values include the hardware energy usage during transmission, reception, and collation of data as adopted in Qureshi et al. (2013); Vipin Pala, Yogitab, Girdhari Singh, and R. P. Yadav (2015); and Moazam and Seyed (2016).

The radio model is adopted in order to analyze and compare the developed model with that of Samayveer et al. (2016). The radio model's energy dissipation values include the hardware energy usage during transmission, reception, and collation of data, as adopted in Qureshi et al. (2013); Vipin Pala, Yogitab, Girdhari Singh, and R. P. Yadav (2015); and Moazam and Seyed (2016).

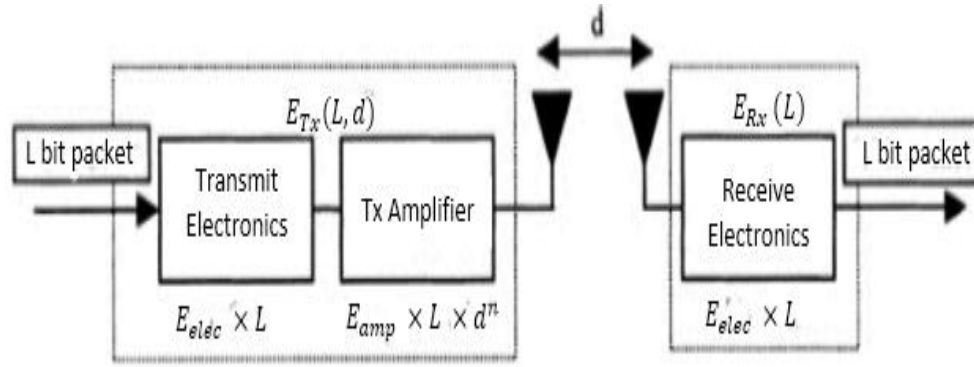


Fig. 2.9 Adopted radio mode (Monica and Dipak 2014)

Let E_{elec} be the energy consumed per bit to run the receiver (E_{RX}) or transmitter (E_{TX}) circuits of the nodes. E_{fs} is the loss suffered due to transmitting the data through free space, and E_{mp} is the loss suffered due to transmitting the data through a multi-path depending on the radio model in question. d is the distance from the sending node to the receiving node. The free space model is used when the distance between the receiver and the transmitter is less than the threshold. The multi-path fading channel model is considered if the distance between the receiver and the transmitter is greater than the threshold distance, which considers two-ray ground propagation. D_o is the threshold distance.

Hence, in order to achieve an acceptable Signal-to-Noise Ratio (SNR) in transmitting an 1-bit message over a distance d , the energy expended by the radio is given by:

$$E_{Tx}(L, d) = \begin{cases} L \cdot E_{elec} + L \cdot \epsilon_{fs} d^2, & \text{if } d < d_o \\ L \cdot E_{elec} + L \cdot \epsilon_{mp} d^4, & \text{if } d \geq d_o \end{cases} \quad (2.8)$$

Where E_{elec} is the energy used per bit (L) to run the receiver or transmitter circuit of the node. ϵ_{fs} represents the free space model used if the distance is less than the threshold d_o , and ϵ_{mp} represents the multi-path model used if the distance is greater than or equal to the threshold d_o .

Thus, the threshold distance d_o is given by:

$$d_o = \sqrt{\frac{E_{fs}}{E_{mp}}} \quad (2.9)$$

Equation (2.9) will be adopted for the value of d_o as in C. Divya, N. Krishnan, and A. Petchiammal (2013); Monica and Dipak (2014); Shilpa et al. (2014); Neeraj Kumar, Sudhanshu Tyagi, and Der-Jiunn Deng (2014); and Moazam and Seyed (2016).

Energy dissipation of each cluster member is given by (2.10) as in (Zhen et al, 2013; Shilpa et al, 2014; Lawal, 2017 and Kalu, 2017):

$$E_{non-CHs(adopted)} = L[E_{elec} + E_{fs} \times (minDisCH)^2] \quad (2.10)$$

where, $E_{non-CHs(adopted)}$ is the energy dissipated by all the non-cluster heads nodes in each cluster, E_{elec} be the energy dissipated per bit to run transmitter (E_{TX}) or receiver (E_{RX}) circuits of the nodes, E_{fs} is free space loss and $minDisCH$ is the minimum distance between the cluster head and each member of the cluster.

While the energy dissipation by the cluster head during a round $E_{CH(adopted)}$ is also adopted from Zhen et al, 2013; Shilpa et al, 2014; Lawal, 2017 and Kalu, 2017, is as given in (2.11).

$$E_{CH(adopted)} = L \left[\left(\frac{N}{K} - 1 \right) E_{elec} + \frac{N}{K} E_{DA} + E_{elec} + E_{fs} \times (d_{CB})^2 \right] \quad (2.11)$$

The first part of (2.11) shows the energy dissipated, E_{elec} by cluster head node to receive $\left(\frac{N}{K} - 1 \right)$ messages from sensor nodes associated with it. Next is energy dissipated in data aggregation E_{DA} .

The energy used in transmission of data to the sink is considered as follows. When the distance between the Cluster Head (CH) and the sink is less than the threshold d_o , transmission occurs in free space. Here, d_{CB} is the distance between the CH and the Base Station (sink), while L is the data bits received from each node within the cluster. (2.11) can be simplified as (2.12):

$$E_{CH(adopted)} = L \left[\frac{N}{K} E_{elec} + \frac{N}{K} E_{DA} + E_{fs} \times (d_{CB})^2 \right] \quad (2.12)$$

However, when the distance $d_{CB} \geq d_o$ then we consider the transmission between the CH to sink to be a case of multi-path and (2.12) becomes.

$$E_{CH(adopted)} = L \left[\frac{N}{K} E_{elec} + \frac{N}{K} E_{DA} + E_{mp} \times (d_{CB})^4 \right] \quad (2.13)$$

Hence, the adopted mathematical model for the total energy dissipation of a cluster (E_C) during transmission for the adopted is given as the sum of the energy dissipated by the CH ($E_{CH(adopted)}$) and the energy dissipated by the cluster members ($E_{non-CHs(adopted)}$) given in (2.14):

$$E_C = E_{CH(adopted)} + E_{non-CHs(adopted)} \quad (2.14)$$

Hence, putting (2.12) and (2.10) into (2.14) gives (2.15).

$$E_C = L \left[\frac{N}{K} E_{elec} + \frac{N}{K} E_{DA} + E_{fs} \times (d_{CB})^2 \right] + L [E_{elec} + E_{fs} \times (minDisCH)^2] \quad (2.15)$$

If the base station is situated at a distance greater than or equal to the threshold d_o , that is when $d_{CB} \geq d_o$ then (2.15)

becomes:

$$E_C = L \left[\frac{N}{K} E_{elec} + \frac{N}{K} E_{DA} + E_{mp} \times (d_{CB})^4 \right] + L [E_{elec} + E_{fs} \times (minDisCH)^2] \quad (2.16)$$

3.0 Methodology

3.1 Introduction

This section presents a detailed methodology and the modified mathematical model on how WSNs can be made to consume less energy when deployed to an environment to be monitored, for optimum operation and lasting services.

3.2 Methodology / Protocol Details

The total number of nodes in this research is 100. Out of these, 10 nodes have more energy than the others, i.e., each of the 10 nodes has 1.6 times more energy than any of the remaining 90 sensor nodes. The formers are termed as *Super Nodes*, while the latter are considered to be *Normal Nodes*. The network area is a 100 m × 100 m field.

The protocol operation is divided into two parts: Cluster Head (CH) selection and Cluster Formation. The cluster formation part is the main concern of this work. Initially, all the nodes are randomly deployed in the field to be monitored, with the sink located at the center of the network field.

The sink sends a beacon message to all the nodes in the network. All nodes respond by sending their ID and location to the sink/base station (BS). The sink uses this information to compute:

- The distances between each node and the sink using Equation (1.1).
- The distances between each node and the others using Equation (1.2).

This information is then transmitted back to the nodes.

The nodes create a Lookup Table (LT) and store the distances relevant to them. The sink then generates random numbers between 0 and 1 and sends these generated numbers to the nodes. Each node chooses a number from the generated set. If the number chosen by the node is less than the threshold $T(s)$ as in Equation (2.3), and the sensor node belongs to set G, it is eligible to become a CH; otherwise, it is not. Note that G is the set of all nodes that did not become CHs in the previous round.

The sink also sets up a TDMA schedule and transmits this schedule to the selected CH nodes in the network. The TDMA schedule ensures that there are no collisions among the data packets sent by the CHs to the base station. Once the TDMA schedule is known by all CHs in the network, they must store it in their LT and wait to communicate with the sink only in their allocated time slots.

Every CH then sends a cluster formation request message, along with its distance to the base station, to all the nodes within its communication range in the network.

3.2.1 Cluster Formation Process

With reference to Fig. 3.1, the cluster formation process was determined by considering the following cases, assuming that d_{NC} is the distance from the node to the Cluster Head (CH) and d_{CB} is the distance between the CH and the Base Station (BS).

Case One: When the distance d_{NC} is the shortest to the cluster head

From Fig. 3.1, it can be observed that the node in question has the shortest route to CH2 compared to other CHs. However, it can also be established that the overall aggregate distance in this case is the longest to the Base Station. This implies that if the node chooses to route its sensed data through CH2, more energy will be dissipated. Such a situation is undesirable as it leads to high energy consumption in the network.

This scenario occurs in schemes that rely on minimum distance or the strongest RSS as their cluster formation criterion. The new scheme presented in this work addresses this limitation by allowing the node to use the aggregate distance to the BS in choosing the appropriate CH to send its data. This ensures that the data does not traverse a longer path to the BS, thereby saving vital energy and decreasing delay. Consequently, the problem of extra communication is reduced or eliminated.

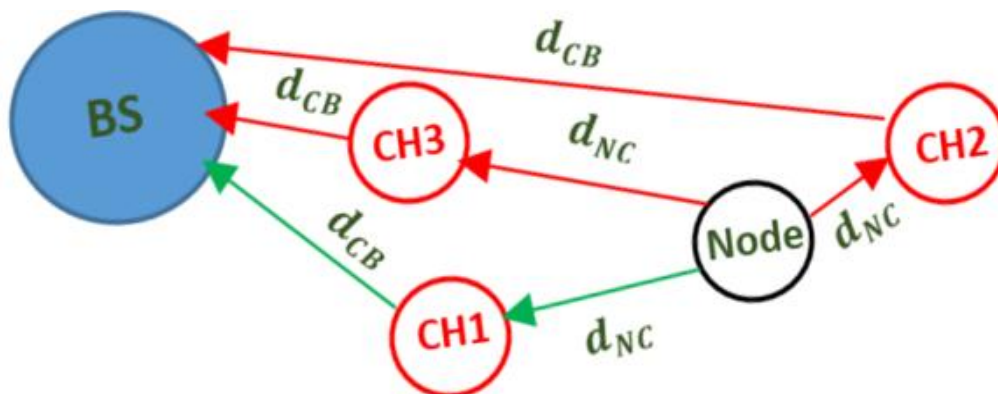


fig. 3.1: Cluster formation process

Case Two: When the distance d_{NC} is greater than the distance d_{CB}

Observing Fig. 3.1, it can be strongly established that the route through CH3 is the shortest route to the BS from the non-CH node in question. However, in this case, the route from the non-CH node to CH3 (d_{NC}) is longer than the route from CH3 to the BS (d_{CB}). This means that the non-CH node transmits its data through a longer distance to CH3 than the distance CH3 has to transmit its aggregated data to the BS.

In this situation, the non-CH node depletes more energy than even CH3, which is undesirable in normal WSN operation. One of the essences of clustering in WSNs is to allow only a few nodes (CHs) to communicate with the BS through longer distances. Hence, only the CHs are “punished” with communicating their aggregated data via longer distances. Therefore, if a cluster member (CM) depletes more energy in transferring data to the CH than the CH does in transferring aggregated data to the BS, that CM will drain its precious energy faster than even the CH. This will lead to premature collapse of the network and a decrease in its lifespan.

Case Three: When the distance d_{CB} is greater than the distance d_{NC}

Lastly, when the distance d_{NC} is less than d_{CB} and the aggregate distance (d_{agg}) is the shortest route to the BS without punishment, then it provides a perfect route to transmit the sensed data from the non-CH node through CH1 to the BS. Under this condition, the energy dissipated by the network will be less compared to the other two cases explained above.

As observed from various research works, the distance through which the data travels is directly proportional to the rate at which energy is dissipated in a wireless sensor network. This makes the CHs dissipate more energy than each of their members. Hence, transmission of aggregated data from the CH to the BS requires significant energy sacrifice.

In this scheme, it is clearly noted that the CMs are “punished” only slightly so that the overall energy of the network is saved by transmitting their sensed data through the shortest route without unnecessary punishment to the BS. In this way, the network consumes less energy.

When these conditions are met, the node accepts to be a CM of that particular CH. The CM then senses the required data and communicates it to its CH for fusion and aggregation. The CH performs fusion and aggregation on all the data sent by its CMs to avoid redundancy and transmits it directly to the BS. Subsequently, the whole process repeats until the total energy of the network is depleted. At any point in time, if a node depletes all its energy, it is assumed to be dead.

3.3 Modification of the Adopted Mathematical Model for the Network Energy Consumption

As can be seen from Equation (2.1), distance is the parameter that directly affects the rate at which energy is consumed. In other words, the longer the distance travelled by the data, the more energy is consumed by the network. Thus, the energy consumed by the network is directly proportional to the distance travelled by the data, right from the non-cluster head sensor nodes via the CH to the sink.

Accordingly, the mathematical model of the energy dissipated by a non-cluster head node, as adopted from Equation (2.10), is modified to become (3.1):

$$E_{non-CHs(modified)} = L[E_{elec} + E_{fs} \times (distance1)^2] \quad (3.1)$$

where:

where, $E_{non-CHs(modified)}$ is the modified energy dissipated by each the non-cluster head node in each cluster, E_{elec} be the energy dissipated per bit to run transmitter (E_{TX}) or receiver (E_{RX}) circuits of the nodes. E_{fs} is free space loss. L is the number of bit transmitted and $distance1$ is the distance between the CH and each member of the cluster. Note that $distance1$ is not equal to the minimum distance to CH.

Equation (3.1) therefore modifies the adopted model to account for the aggregate distance travelled by the data, ensuring that both the intra-cluster transmission (d_{NC}) and the inter-cluster transmission (d_{CB}) are considered in the overall energy consumption.

Therefore, (2.14) now becomes:

$$E_C = E_{CH} + E_{non-CHs(modified)} \quad (3.2)$$

That is

$$E_C = L \left[\frac{N}{K} E_{elec} + \frac{N}{K} E_{DA} + E_{fs} \times (distance)^2 \right] + L[E_{elec} + E_{fs} \times (distance1)^2] \quad (3.3)$$

where E_C is the total energy consumed by the cluster, N is the total number of nodes in the network, K is the number of CHs in the network, E_{DA} is the energy consumed due to data aggregation E_{elec} be the energy dissipated per bit to run transmitter (E_{TX}) or receiver (E_{RX}) circuits of the nodes. E_{fs} is free space loss. L is the number of bit transmitted and $distance1$ is the distance between the CH and each member of the cluster while distance is the distance between the CH and the BS. Note that $distance1$ is not equal to the minimum distance to CH.

If the signal undergoes a multipath that is the Base station is situated at a distance greater than or equal the threshold d_o , then (3.3) becomes:

$$E_C = L \left[\frac{N}{K} E_{elec} + \frac{N}{K} E_{DA} + E_{mp} \times (distance)^4 \right] + L[E_{elec} + E_{fs} \times (distance1)^2] \quad (3.4)$$

Note that $distance1$ is not equal to the minimum distance between the CM and CH. While E_{mp} is multipath loss.

(3.1), (3.2), (3.3) and (3.4) are the modified mathematical model of (2.10), (2.14), (2.15) and (2.16) respectively, which aid in saving the overall energy of the network.

4.1 Simulation Results and Discussion

The results are established based on the network setup variables displayed in Table 4.1. The performance of the proposed IAACF protocol is analyzed in comparison to hetDEEC-3 (Samayveer et al., 2016) through performance metric of Lifetime of the network, this metric was evaluated after carrying out a series of rounds data transfer.

Table 4.1 Network Parameters

| Network Parameters | Value |
|--|-----------------------------|
| Network size | 100 × 100 |
| Packet size | 4000bits |
| E_{elect} (Energy consumed in the electronic circuit to transmit or receive the signal) | 50nJ/bits |
| Data aggregation/fusion energy consumption | 5nJ/bit/signal |
| E_{fs} (Energy consumed by the amplifier to transmit at a short distance) | 10nJ/bit/m ² |
| E_{mp} (Energy consumed by the amplifier to transmit at a long distance) | 0.0013pJ/bit/m ⁴ |
| Initial Energy of the sensor nodes used in this work for comparison with hetDEEC-3 of Samayveer et al, 2016. | |
| Normal node | 0.944J |
| Super node | 1.6 × 0.944 ≈ 1.5J |

4.2 Comparison Between IAACF and hetDEEC-3 in Terms of the Total Energy Consumption of the Network

The performance metrics used to compare this work with hetDEEC-3 include network energy consumption, which measures the total energy dissipation of the network. This value is calculated at each round of the protocol. Lower energy dissipation corresponds to a longer lifetime of the network.

The performance of IAACF for WSN in comparison with hetDEEC-3 was evaluated in percentage with respect to the above-mentioned metric using Equation (4.1):

In any WSN, the lifetime or lifespan of the network depends solely on the rate at which energy is consumed in the network per unit round. The total energy consumed by a cluster in the network is given by Equation (3.3). If the Base Station is situated at a distance greater than or equal to the threshold d_0 from the CH, then the total energy consumed by a cluster in the network is given by Equation (3.4) in a round.

Therefore, the total amount of energy consumed by each cluster in the network was deducted from the total energy of the network (100 J) at the end of each round. Table 4.2 shows the amount of energy consumed by IAACF as compared to hetDEEC-3 at various round intervals.

Table 4.2: Comparison between IAACF and hetDEEC-3 in terms of energy consumed in the network

| SUM OF THE ENERGY DISSIPATED BY THE NETWORK VESUS NO. OF ROUNDS | | |
|---|-----------|-------|
| Rounds | hetDEEC-3 | IAACF |
| 0 | 0 | 0 |
| 500 | 16 | 10 |
| 1000 | 34 | 23 |
| 1500 | 52 | 40 |
| 2000 | 68 | 58 |
| 2500 | 80 | 72 |
| 3000 | 90 | 83 |
| 3500 | 96 | 89.8 |
| 4000 | 98 | 93 |
| 4500 | 100 | 95.5 |

Hence, Fig. 4.2 shows the graphical representation of the total energy consumed versus the number of rounds for both hetDEEC-3 and IAACF formed from table 4.4. The total energy consumed in the network was plotted on the Y axis against the number rounds on the X axis. From the Fig. 4.2, it can clearly be seen that, the total energy consumption of IAACF was lower in every round as compared to hetDEEC-3. The energy was completely exhausted in hetDEEC-3 in 4404 rounds. While, the energy of IAACF reach up to 5924 rounds before it was used up.

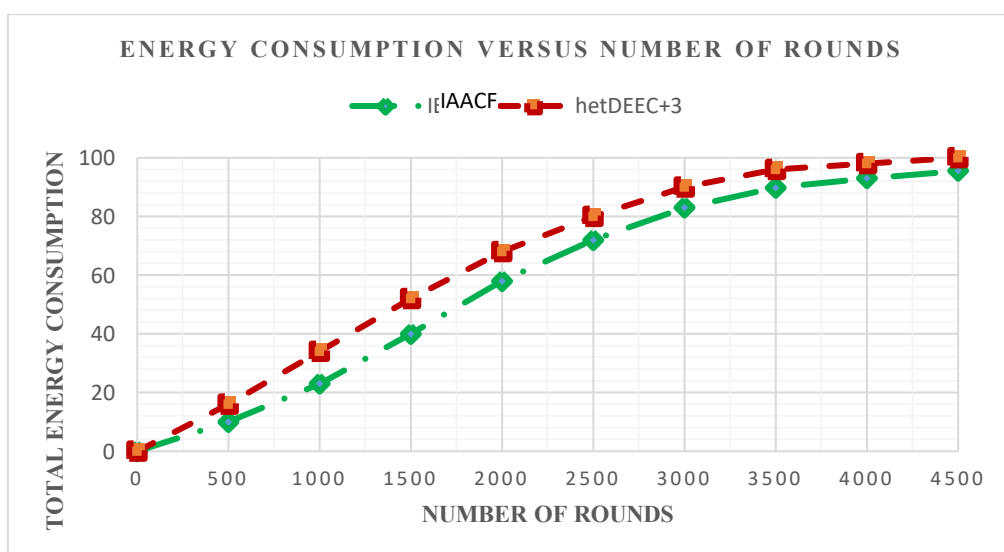


Fig. 4.2 Comparison between IAACF and hetDEEC-3 in terms of energy dissipated in the Network

Performance Evaluation

Similarly, from table 4.2 the percentage improvement at various rounds intervals can be computed using (4.1) as follows:

$$\text{At 1000 rounds interval } performance = \frac{23-34}{23} \times 100 = 47.48\%$$

$$\text{At 2000 rounds interval } performance = \frac{58-68}{58} \times 100 = 17.24\%$$

$$\text{At 3000 rounds interval } performance = \frac{83-90}{83} \times 100 = 8.43\%$$

And the average percentage improvement can be computed using (4.2) as follows.

$$\begin{aligned} \text{Average performance} \\ = \frac{\text{performance at (1000rounds + 2000rounds + 3000rounds)}}{3} \end{aligned} \quad (4.2)$$

$$\therefore \text{Average performance} = \frac{(8.43 + 17.24 + 47.48)}{3} = 24.38\%$$

Based on the results obtained, it is observed that the IAACF reduces energy consumption per round in the network, thereby increasing the lifespan of the network by an average of 24.38%. This indicates a constant energy depletion rate in the sensor nodes, resulting from effective cluster formation in the IAACF scheme.

5.0 COCLUSION

The IAACF for WSN (new) scheme of cluster formation has been presented and evaluated. The scheme applies an aggregate distance from cluster member nodes through the CH to the Base Station for cluster formation. All non-cluster head nodes employ the aggregate distance criterion to select an appropriate CH, thereby improving WSN performance by reducing the rate of energy consumption. Simulation results confirm that the IAACF scheme achieves significant energy savings compared to minimum distance or strongest RSSI criteria used in related literature. Performance analysis shows improved parameters including number of dead sensor nodes per rounds, packets delivered to the sink per rounds, and overall energy content of the system. The IAACF scheme demonstrates a 25.66% performance increase over the hetDEEC-3 scheme based on network lifetime. This work contributes to the development of IAACF for WSN and suggests potential implementation using artificial intelligence techniques such as GA and Fuzzy Logic.

Sponsor

This research was sponsored by the Tertiary Education Trust Fund (TETfund) under the Institution-Based Research (IBR) 2025 intervention.

REFERENCES

1. Adamu, M. L. (2017). *Development of an improved hybrid low energy adaptive clustering hierarchy (LEACH) for wireless sensor network* [Master's dissertation, Ahmadu Bello University, Zaria].
2. Akila, E., & Baby, D. (2016). An improved cluster head selection technique for wireless sensor network using modified genetic algorithm. *The IIOAB Journal*, 7(9), 446–456.
3. Alaa, E. M., Abd Elrazik, S. M., El-Bakry, H. M., Hasan, S. Q., Hasan, A. Q., & Zaid, S. (2016). Challenges in wireless sensor networks. *International Journal of Advanced Research in Computer Science & Technology*, 4(4), 22–27.
4. Paul, A. K., & Sato, T. (2017). Localization in wireless sensor networks: A survey on algorithms, measurement techniques, applications, and challenges. *Journal of Sensor and Actuator Networks*, 6(4), 24. <https://doi.org/10.3390/jsan6040024>
5. Abasi, A., & Sajedi, H. (2016). Fuzzy-clustering based data gathering in wireless sensor network. *International Journal on Soft Computing*, 7(1), 1–10. <https://doi.org/10.5121/ijsc.2016.71011>
6. Asha, & Vineeta. (2013). An election of vice cluster selection approach to improve V-LEACH protocol in wireless network. In *Proceedings of the IEEE Third International Conference on Advanced Computing & Communication Technologies*.
7. Divya, C., Krishnan, N., & Petchiammal, A. (2013). Increase the alive nodes based on the cluster head selection algorithm for heterogeneous wireless sensor networks. *Global Journal of Computer Science and Technology: Network, Web & Security*, 13(9).
8. Fateh, B., Fatima, T., & Khawla, M. (2015). A hybrid approach to extend the lifetime of heterogeneous wireless sensor networks. *Procedia Computer Science*, 63, 136–141.
9. Guangjie, H., Xu, J., Aihua, Q., Rodrigues, J. J. P. C., & Long, C. (2014). A comparative study of routing protocols of heterogeneous wireless sensor networks. *The Scientific World Journal*, 2014, 1–10.

10. Heinzelman, W. R., Chandrakasan, A., & Balakrishnan, H. (2000). Energy-efficient communication protocol for wireless microsensor networks. In *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences* (pp. 1–10). IEEE.
11. Javaid, N., Qureshi, T. N., Khan, A. H., Iqbal, A., Akhtar, E., & Ishfaq, M. (2013). EDDEEC: Enhanced developed distributed energy-efficient clustering for heterogeneous wireless sensor networks. In *International Workshop on Body Area Sensor Networks (BASNet-2013)*.
12. Jha, V., & Sharma, R. (2022). Energy-efficient clustering in heterogeneous wireless sensor networks using RSS. *Sensors*, 22(14), 5123. <https://doi.org/10.3390/s22145123>
13. Kalu, J. C. (2017). *Development of an energy efficient algorithm to prolong the lifetime of wireless sensor networks* [Master's dissertation, Bayero University, Kano].
14. Kannan, G., & Sree, R. R. T. (2015). Energy-efficient distributed cluster head scheduling scheme for two-tiered wireless sensor network. *Egyptian Informatics Journal*, 16(2), 167–174.
15. Kiran, D. N. R., & Dethe, C. G. (2015). Energy-efficient data aggregation in sensor networks with optimized cluster head selection. *International Journal of Computer, Electrical, Automation, Control and Information Engineering*, 9(6).
16. Kumar, D., Aseri, T. C., & Patel, R. B. (2011). A novel multihop energy efficient heterogeneous clustered scheme for wireless sensor networks. *Tamkang Journal of Science and Engineering*, 14(4), 359–368.
17. Lekhi, S., & Singh, S. (2021). LEACH-based energy efficient clustering protocol for wireless sensor networks. *IEEE Access*, 9, 123456–123467. <https://doi.org/10.1109/ACCESS.2021.xxxxxx>
18. Lewandowski, M., & Płaczek, B. (2025). A cluster head selection algorithm for extending last node lifetime in wireless sensor networks. *Sensors*, 25(11), 3466. <https://doi.org/10.3390/s25113466>
19. Lonkar, B., Kuthe, A., Charde, P., Dehankar, A., & Kolte, R. (2024). Hybrid energy-saving cluster head selection for wireless sensor networks. *Peer-to-Peer Networking and Applications*, 18, Article 182. <https://doi.org/10.1007/s12083-025-02002y>
20. Manpreet, K., & Abhilasha. (2014). Survey on variants of distributed energy efficient clustering protocols in heterogeneous wireless sensor networks. *International Journal of Computer Engineering and Applications*, 6(2).
21. Meenakshi, D., & Sushil, K. (2012). An energy efficient level based clustering routing protocol for wireless sensor networks. *International Journal of Advanced Smart Sensor Network Systems*, 2(2).
22. Milica, D., et al. (2012). A survey of military applications of wireless sensor networks. In *Proceedings of the 2012 Mediterranean Conference on Embedded Computing (MECO 2012)* (pp. 196–199).
23. Moazam, B., & Seyed, R. K. T. (2016). Efficient fuzzy logic-based clustering algorithm for wireless sensor networks. *International Journal of Grid and Distributed Computing*, 9(5), 79–88.
24. Farsi, M., Elhosseini, M. A., Badawy, M., Ali, H. A., & Zain Eldin, H. (2019). Deployment techniques in wireless sensor networks, coverage, and connectivity: A survey. *IEEE Access*, 7, 28940–28954. <https://doi.org/10.1109/ACCESS.2019.2902072>
25. Monika, R., & Dipak, W. W. (2014). An energy saving algorithm to prolong the lifetime of wireless sensor network. *International Journal of Wireless & Mobile Networks*, 6(5).
26. Muhammad, A., Ehsan, U. M., Mustafa, R. M., & Xiaopeng, H. (2016). Adaptive energy-efficient clustering path planning routing protocols for heterogeneous wireless sensor networks. *Sustainable Computing: Informatics and Systems*, 10, 1–12.
27. Nadeem, J., Muhammad, B. R., Muhammad, I., Mohsen, G., Zahoor, A. K., Turki, A. A., & Manzoor, I. (2015). An energy-efficient distributed clustering algorithm for heterogeneous WSNs. *EURASIP Journal on Wireless Communications and Networking*, 2015(1), 1–11.
28. Neeraj, K., Sudhanshu, T., & Der-Jiunn, D. (2014). LAEEHSC: Learning automata-based energy efficient heterogeneous selective clustering for wireless sensor networks. *Journal of Network and Computer Applications*, 39, 111–124.
29. Prachi, U., & Preeti, S. (2014). Review of heterogeneous/homogeneous wireless sensor networks and intrusion detection system techniques. In *Proceedings of the International Conference on Recent Trends in Information, Telecommunication and Computing (ITC)*.
30. Qureshi, T. N., Javaid, N., Khan, A. H., Iqbal, A., Akhtar, E., & Ishfaq, M. (2013). BEENISH: Balanced energy efficient network integrated super heterogeneous protocol for wireless sensor networks. In *International Workshop on Body Area Sensor Networks (BASNet-2013)*.
31. Ravneet, K., Deepika, S., & Navdeep, K. (2013). Comparative analysis of LEACH and its descendant protocols in wireless sensor network. *International Journal of P2P Network Trends and Technology*, 3(1).
32. Al-Omari, S. A. K., & Sumari, P. (2010). An overview of mobile ad hoc networks for existing protocols and applications. *Journal of Applications of Graph Theory in Wireless Ad hoc Networks and Sensor Networks*, 2(1).
33. Singh, S. (2016). Energy-efficient multilevel network model for heterogeneous WSNs. *Engineering Science and Technology, an International Journal*, 19(1), 1–10.
34. Singh, S., Aruna, M., & Rajeev, K. (2016). Energy efficient heterogeneous DEEC protocol for enhancing lifetime in WSNs. *Engineering Science and Technology, an International Journal*, 19(2), 1–8.

35. Mohamed, S. A. E. (2011). Why the accuracy of the received signal strengths as a positioning technique was not accurate? *International Journal of Wireless & Mobile Networks*, 3(3), 69–82. <https://doi.org/10.5121/ijwmn.2011.3306>
36. Sanay, A., & Nima, J. N. (2016). Deployment strategies in the wireless sensor network: A comprehensive review. *Computer Communications*, 76, 1–12.
37. Sandeep, S. M., & Mamta, K. (2014). Efficient cluster head selection scheme for wireless sensor network using deterministic protocol. *IOSR Journal of Computer Engineering*, 16(3), 62–67.
38. Shilpa, M., Jyoteesh, M., & Sandeep, S. (2014). An energy balanced QoS-based cluster head selection strategy for WSN. *Faculty of Computers and Information, Cairo University, 2014*, 1–8.
39. Shipra, S., & Karamjot, K. (2016). Comparative analysis of homogeneous and heterogeneous protocols in WSN. *International Journal of Science and Research*, 5(6).
40. Kole, S., Vhatkar, K. N., & Bag, V. V. (2014). Distance-based cluster formation technique for LEACH protocol in wireless sensor network. *International Journal of Application or Innovation in Engineering & Management*, 3(3), 1–7.
41. Srividhya, V., & Shankar, T. (2019). Energy reckoning distance-based clustering for spectrum aware cognitive radio wireless sensor networks. *International Journal of Innovative Technology and Exploring Engineering*, 9(2).
42. Sudeep, T., Neeraj, K., & Rodrigues, J. J. P. C. (2015). A systematic review on heterogeneous routing protocols for wireless sensor network. *Journal of Network and Computer Applications*, 55, 123–142.
43. Das, S. (2014). *Design and development of energy-efficient routing protocol for heterogeneous wireless sensor networks* [Master's dissertation, Jadavpur University, Kolkata].
44. Samant, T., Mukherjee, P., Mukherjee, A., & Datta, A. (2017). TEEN-V: A solution for intra-cluster cooperative communication in wireless sensor network. In *International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC 2017)* (pp. 209–213).
45. Thiriveni, G. V., & Ramakrishnan, M. (2016). Distributed clustering-based energy efficient routing algorithm for heterogeneous wireless sensor networks. *Indian Journal of Science and Technology*, 9(3).
46. Tuba, F., & Meenakshi, Y. (2016). Modified clustering algorithm for energy efficiency utilizing fuzzy logic in WSN (MCF). *International Journal of Computer Science and Application*, 5(2).
47. Chauhan, V., & Soni, S. (2019). Load-balanced energy efficient cluster-chain-based hybrid protocol for wireless sensor networks. *International Journal of Recent Technology and Engineering*, 8(3), 3561–3569. <https://doi.org/10.35940/ijrte.C5284.098319>
48. Vipin, P., Yogita, B., Girdhari, S., & Yadav, R. P. (2015). Effect of heterogeneous nodes location on the performance of clustering algorithms for wireless sensor networks. *Procedia Computer Science*, 57, 1042–1048.
49. Vivek, M., & Catherine, R. (2004). Homogeneous vs heterogeneous clustered sensor networks: A comparative study. In *Proceedings of a Conference on Sensor Networks* (pp. 1–6).
50. Ahmad Najib, Y. N., Daud, H., & Abd Aziz, A. (2020). Singular value thresholding algorithm for wireless sensor network localization. *Mathematics*, 8(3), 437. <https://doi.org/10.3390/math8030437>
51. Zhanyang, X., Yue, Y., & Jin, W. (2013). An energy-efficient clustering routing algorithm for heterogeneous wireless sensor networks. *NGCIT 2013, ASTL*, 27, 85–93.
52. Zhen, H., Li, Y., & Gui-Jun, Z. (2013). Efficient and dynamic clustering scheme for heterogeneous multi-level wireless sensor networks. *Acta Automatica Sinica*, 39(4).

CITATION

Yako, A. A., & Shuaibu, D. S. (2026). "Improved Algorithmic Approach to Cluster Formation in Wireless Sensor Networks" (IAACF). In *Global Journal of Research in Engineering & Computer Sciences* (Vol. 6, Number 1, pp. 60–73). <https://doi.org/10.5281/zenodo.18358564>