PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA
Ministry of Higher Education and Scientific Research
Mohamed Khider University – BISKRA
Faculty of Exact Sciences, Natural and Life Sciences
Department of Computer Science

N d'ordre:RTIC7/RTIC/M2/2022

THESIS

Presented for the Academic Master's degree in **Computer Science**

Pathway : Information and Communication Technologies and Networks(RTIC)

By : Djehiche Aicha

Title

Study of Clustering techniques in Wireless Sensor Networks(WSNs) (LEACH and K-means)

Presented on the 19 /06 /2023 board of the jury composed of:

Mme.	Sahraoui Somia	MCA	President
Mme.	Aloui Imane	MCB	Supervisore
M.	Laanani Sadek	MCD	Examiner

Academic year 2022-2023

Acknowledgement

In the name of God, the Most Merciful, and may prayers and peace be upon the most honorable of messengers, Muhammad El-Amin, peace and peace be upon him.

First, I want to thank God for everything He has given me and all the strength He has given me. I would like to express my sincere acknowledgment to all those who contributed to the completion of this thesis.

First and foremost, I would like to thank my thesis advisor, [Dr.Aloui Imene], for her support, insightful guidance, and availability throughout this project.

I extend my gratitude to the members of the thesis committee, for agreeing to evaluate this thesis and for their valuable comments and suggestions.

A heartfelt thanks go to my dear mother and my sisters and my brother, and my friend for their unwavering support, encouragement, and unconditional love. their presence and encouragement have been instrumental in helping me overcome challenges and stay focused on my goals.

I want to express my appreciation to my colleagues and classmates who have engaged in enriching discussions and provided support throughout this academic journey.

Abstract

The wireless sensor network (WSN) faces a significant challenge due to the limited energy resources of its sensor nodes. Addressing this issue, the clustering technique in WSNs helps in achieving energy efficiency, this study focuses on modifying one of the most widely used Clustering algorithms for data communication in WSNs: LEACH (Low Energy Adaptive Clustering Hierarchy). The revised version, named "K-means-LEACH," incorporates an intermediate cluster head for efficient data transmission, thereby extending the network's lifetime and enabling the transmission of more data compared to the original protocol.
To evaluate the effectiveness of the proposed algorithm in improving the network's lifetime, MATLAB conducted simulations. The simulation results confirmed that the modified system outperformed the LEACH protocol, enhancing network lifetime.

Keywords: wireless sensor network, Clustering, LEACH, Energy consumption

ملخص

تواجه شبكة الاستشعار اللاسلكية تحديا كبيرا بسبب موارد الطاقة المحدودة لعقد الاستشعار الخاصة بها. لمعالجة هذه المشكلة ، تساعد تقنية التجميع في شبكات في تحقيق كفاءة الطاقة ، وتركز هذه الدراسة على تعديل واحدة من أكثر خوارزميات التجميع استخداما لاتصالات البيانات في شبكات (التسلسل الهرمي للتجميع التكيفي منخفض الطاقة). تتضمن النسخة المعدلة ، رأس مجموعة وسيطة لنقل البيانات بكفاءة ، وبالتالي إطالة عمر الشبكة وتمكين نقل المزيد من البيانات مقارنة بالبروتوكول الأصلي.
لتقييم فعالية الخوارزمية المقترحة في تحسين عمر الشبكة ، أجريت عمليات محاكاة باستخدام ماتلاب أكدت نتائج المحاكاة أن النظام المعدل تفوق على بروتوكول ، مما أدى إلى تحسين عمر الشبكة..

الكلمات المفتاحية: شبكة الاستشعار الاسلكية،التجميع،استهلاك الطاقة،LEACH

Contents

List of Figures

General introduction

Today, the rapid progress in technology has forged a seamless alliance between computer science and electronics. In this symbiotic relationship, the realm of wireless and electronic technology stands out as one of the largest and most influential fields, offering cost-effective solutions with remarkable efficiency. Among the groundbreaking technologies emerging from this synergy, Wireless Sensor Networks (WSNs) hold a position of paramount importance. WSNs have the potential to revolutionize our daily lives by simplifying numerous tasks and activities.

This convergence has paved the way for astounding advancements in wireless communication, mobile networks, and the continuous miniaturization of sensors. With their vast and diverse range of applications, wireless sensors enable remote tracking and precise control of the physical environment, leading to enhanced monitoring accuracy and efficiency.

As the field of wireless sensor networks continues to evolve, researchers and practitioners strive to overcome challenges such as energy efficiency, network scalability, data reliability, and security.

Through innovative solutions, novel communication protocols, and advanced algorithms, WSNs are being fine-tuned to unlock their full potential. The constant exploration of new frontiers and the fusion of WSNs with emerging technologies like the Internet of Things (IoT) and Artificial Intelligence (AI) promise even greater possibilities and widespread adoption in the future.

Clustering in WSN is an effective way to structure the network to facilitate communication and data management.

clustering in WSNs helps in achieving energy efficiency, load balancing, and prolonging network lifetime.

The main objective of this work is to propose an algorithm for wireless sensor networks that address the challenges of energy consumption. In this study, we propose a novel protocol called K-means-LEACH (K-means-Low-Energy Adaptive Clustering Hierarchy), which is a modified version of the LEACH protocol known for its superior power performance compared to other routing protocols.

Building upon the strengths of LEACH and k-means, By combining these two approaches, our method aims strives to strike a balance between energy optimization and the sustainability of the network.

The K-means-LEACH protocol ensures efficient energy distribution and resource allocation among the sensor nodes, thereby improving the network's energy efficiency, Thus extending its life of the network.

Our work is structured as follows:

Chapter 01: Overview Of Wireless Sensor Networks(WSN).

Chapter 02: present the Clustering in Wireless Sensor Networks (WSN).

Chapter 03: Conception of the proposed approach.

Chapter 04:The results of simulations that have been conducted to evaluate the functionality of our approach.

Chapter 1

Overview Of Wireless Sensor Networks

1.1 Introduction

Today, technological advances have created a perfect symbiosis between computer science and electronics. Among them, the field of wireless and electronic technology is regarded as one of the largest and most popular fields, given its low cost and high efficiency.

WSN is one of the most important technologies that will revolutionize the way we live, making many tasks in life easier today. This cohabitation has allowed a dazzling development of technologies in communication through wireless, mobile networks, and constantly miniaturized sensors. Wireless sensors have a very wide and varied field of application, WSN makes it easier to remotely track and control the physical environment with greater accuracy.

1.2 Sensor

1.2.1 Definition

A sensor node is a device that transforms observed physical quantities [9](temperature, sound, chemicals, or the presence of certain objects). Then convert them into binary data that can be used by computer systems. [10] Information can be collected(captured), modified, stored, and transmitted to base stations or other sensors [11].

1.2.2 Architecture of Sensor

1.2.2.1 Hardware architecture of a sensor

A sensor is composed of 4 units:

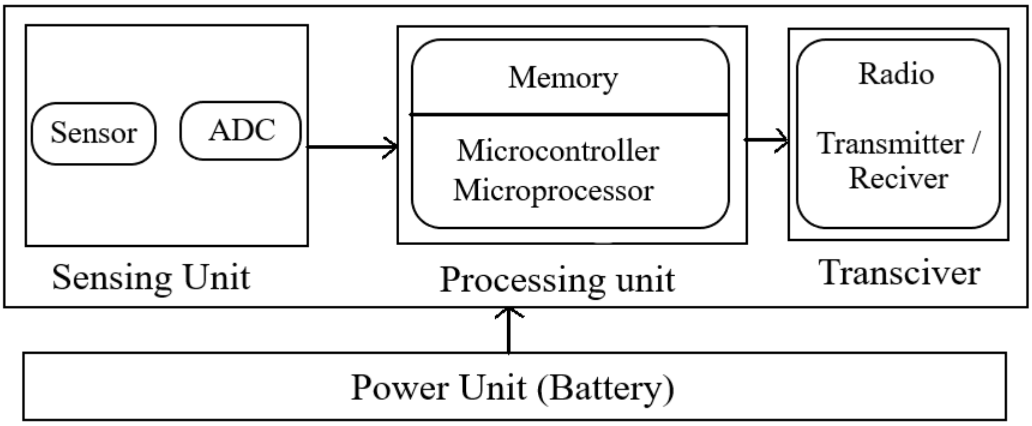


Figure 1.1: Sensor Components [1]

1. **Sensing Unit**
   All sensor devices are equipped with a sensing unit. Usually, he is divided into two sub units. Sensor part and analog-to-digital. The sensor part includes camera, video, sound, and/or scalar sensors and analog-to-digital converters. The software converts the analog signal generated by the sensor node into a digital signal and sends it to the processing unit [12].

2. **Power Unit**
   Power supply units provide power to sensor nodes, and sensors use power in many areas, such as the sensing environment and data processing that comes from sensor nodes and is communicated to other sensor nodes. Many studies show that it consumes more energy than any other process. The basic energy sources for sensor nodes are electrochemical materials such as NiMH, NiZn, and lithium-ion cells. [12].

3. **Communication Unit**
   A communication unit is a subsystem that stabilizes the interface between a device and a network, allowing transmission and reception with the help of communications. software [12].

4. **Processing Unit**
   Upon receiving information or data from a sensor node, the processing unit begins executing it as a coordinated in the system software. It has interacted with storage units and communication tasks [12].

1.3 Wireless Sensor Networks

1.3.1 Definition

is a special type of ad hoc networks [13], consists of a large number of widely distributed devices with sensors [14] sensors deployed in geographical areas called catchments or areas of interest, [15].

Each of these nodes can collect and transfer data to a base station (called a "sink"). The latter takes the information reported by various sensors and sends it via the internet or satellite to a processing center [16]

1.3.2 Architecture of WSN

When deploying many sensor nodes over a large area to monitor the physical environment, it is also important to network these sensor nodes. Sensor nodes within a WSN communicate not only with other sensor nodes, but also with base stations (BSs) using radio communications. [2]

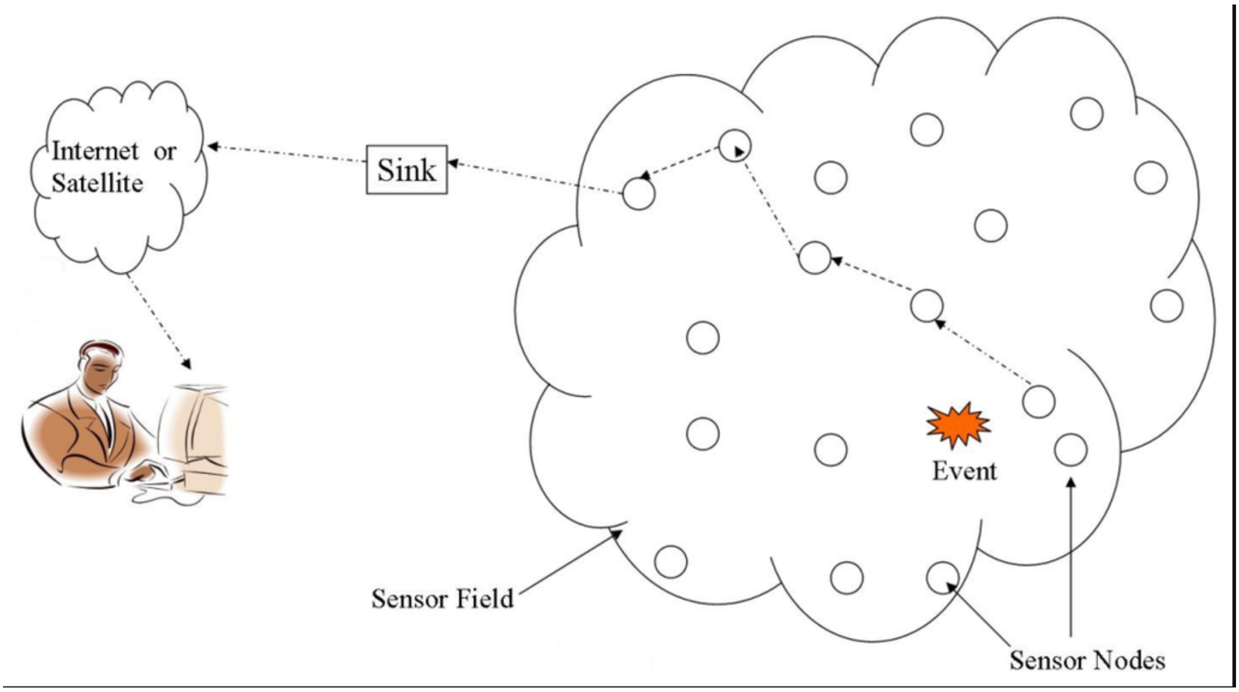


Figure 1.2: WSN Architecture [2]

1.3.3 Wireless sensor networks topologies

1. **Star Topology**

   Star topology consists of a central node, called a sensor network coordinator or sink, and multiple wireless sensor nodes. In this topology, all sensor nodes send data directly to the coordinator. [9] In this topology, a central node takes over control and coordination of sensor nodes that communicate only with the central node. [9]

   (1) **Advantages** This topology is very efficient and relatively simple as it limits data transfer as much as possible. [9]

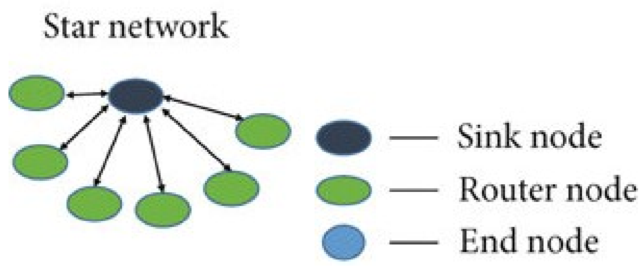


Figure 1.3: Star Topology [3]

2. **Mesh Topology**

   This topology consists of a series of connected nodes. They connect to each other and act as routers. Moreover, all nodes have the same role. Each node determines the best path to transfer data through the network to a well (coordinator or sink) and then sends that data to its neighbors. This feature helps keep the network up and running in the event of a node failure. [17]

   (1) **Advantages** When there is no data transfer, the node goes into sleep mode to save power. [17]

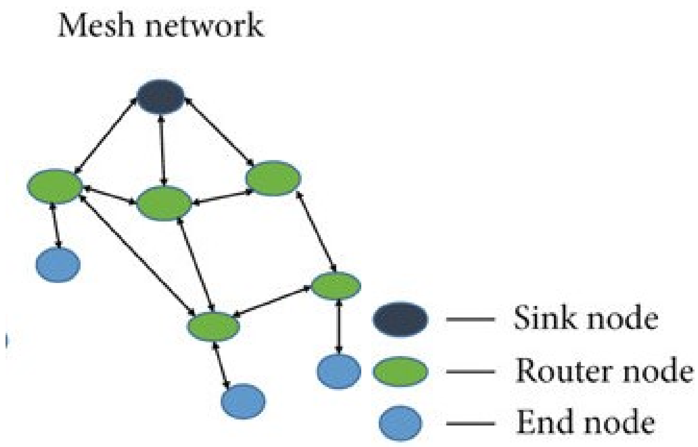


Figure 1.4: Mesh Topology [3]

3. **Tree Topology**

   In terms of the tree topology, all sensors used in the sensor field form a logical tree. By doing so, you have two types of knots. One is the parent node and the other is the leaf node. Data packets are routed from a leaf node to its parent node. On the other hand, a receiving node that receives data from a child node accumulates the data in its own data and then sends the data to the receiving parent node. [18]

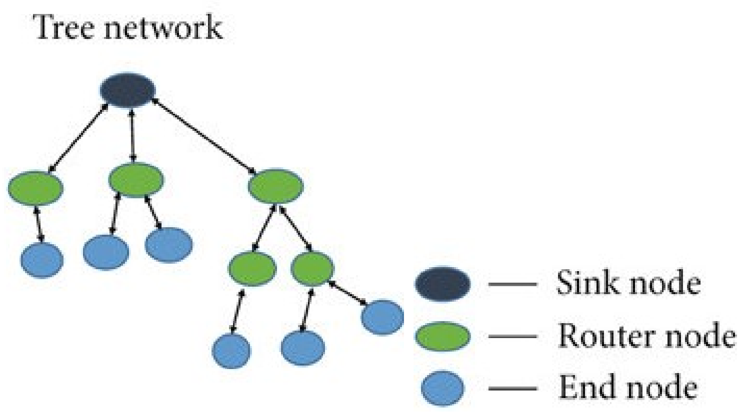


Figure 1.5: Tree Topology [3]

1.3.4 WSN Characteristics

1. **Energy efficient**

   WSN energy is used for various purposes such as computation, communication, and storage. Sensor nodes use more energy than any other node for communication.

When it runs out of energy, it often becomes ineffective because there is no way to recharge it. Therefore, protocol and algorithm development should consider power consumption at the design stage. [19]

2. **Multi-hop communication**

   WSN has a large number of sensor nodes deployed. Therefore, a viable way to communicate with a sink or base station is through the routing path with the help of intermediate nodes. If it needs to communicate with other nodes or base stations that lack radio frequencies, it must do so via multi-hop routes through intermediate nodes. [19]

3. **Dynamic network topology**

   WSNs are generally dynamic networks. The network topology may change frequently due to sensor node failures due to battery depletion, communication path disruptions, and sensor node additions to the network. Therefore, WSN nodes must incorporate reconfiguration and self-adaptation capabilities. [19]

4. **Self-organization**

   Reliable maintenance of these types of networks requires a great deal of self-organization. Network topologies can be unstable, so the network must be able to self-organize in order for the application to continue. [19]

5. **Communication Capabilities**

   WSNs typically communicate using radio waves over radio channels. It has the characteristic of communicating over short distances with a narrow and dynamic bandwidth. Communication channels are either bidirectional or unidirectional. Unattended and hostile operating environments make it difficult for WSNs to run smoothly. Therefore, hardware and software for communication must consider robustness, security, and resilience. [19]

1.3.5 Routing protocols of WSN

Data on the network is routed from one node to another, reaching from source to destination. High network throughput requires the shortest routing path. There are three categories of routing protocols, depending on the network structure: flat, hierarchical, location-based . [1]

1. **Flat based routing** This type of routing category is multi-hop routing where all nodes operate simultaneously. If the sink or base station can request data in its region, all nodes in that region will send that data. As a result, this type of routing is energy consumption. [1]

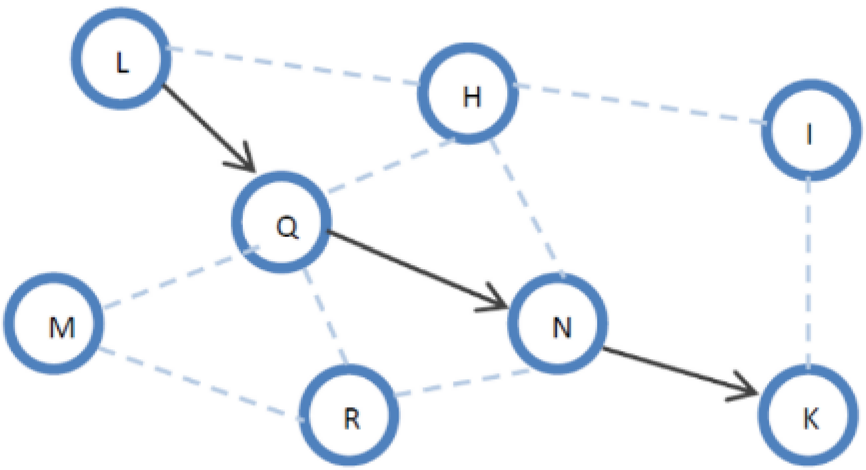


Figure 1.6: Flat based routing [4]

Some flat routing : [20]

(1) **Sensor Protocol for Information via Negotiation (SPIN).**

(2) **Directed Diffusion (DD).**

(3) **Rumor Routing (RR).**

(4) **Gradient-based Routing (GBR).**

(5) **Constrained Anisotropic Diffusion Routing (CADR).**

(6) **Energy Aware Routing (EAR).**

2. **Hierarchical based routing** This approach is based on the formation of clusters (common areas). The principle is to forward the data collected by each node of the

cluster to its zone leader (cluster head). After processing the intersection, the zone leader sends it to the next destination. The advantage is that by minimizing the number of messages circulating in the network, communication and energy costs can be reduced. [4]

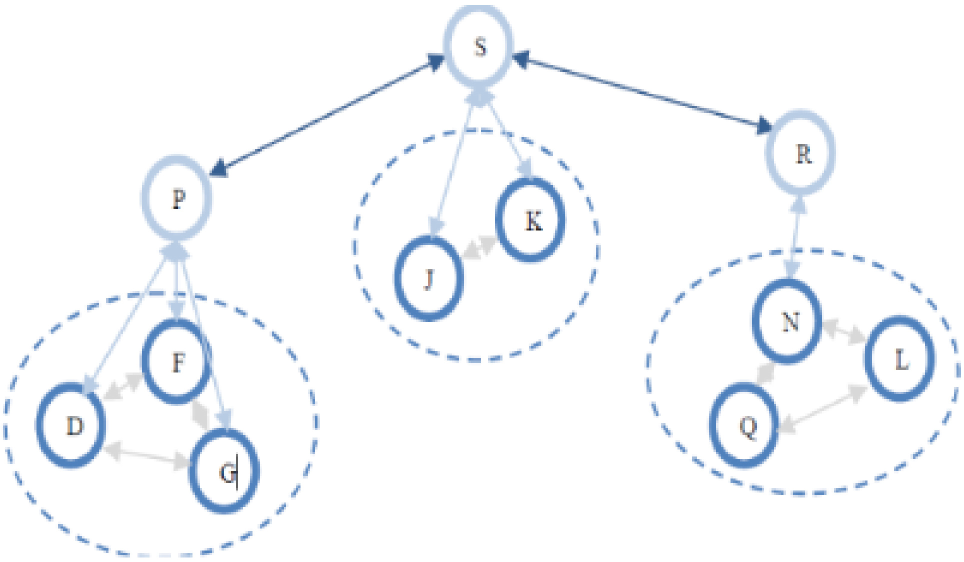


Figure 1.7: Hierarchical based routing [4]

Some hierarchical routing: [20]

(1) **Low energy adaptive clustering hierarchy(LEACH)**

(2) **Minimum energy communication network (MECN).**

(3) **Power-efficient gathering in sensor information systems(PEGASIS)**

(4) **Self-Organizing Protocol (SOP)**

(5) **Threshold Sensitive Energy Efficient Protocols: TEEN and APTEEN**

3. **Location based routing protocols** Determining the geographic location of sensor nodes within a watershed is paramount to the WSN's data routing mechanism. This location information allows us to calculate the sensor's position and distance to create the shortest path between the source node and its destination. [4]

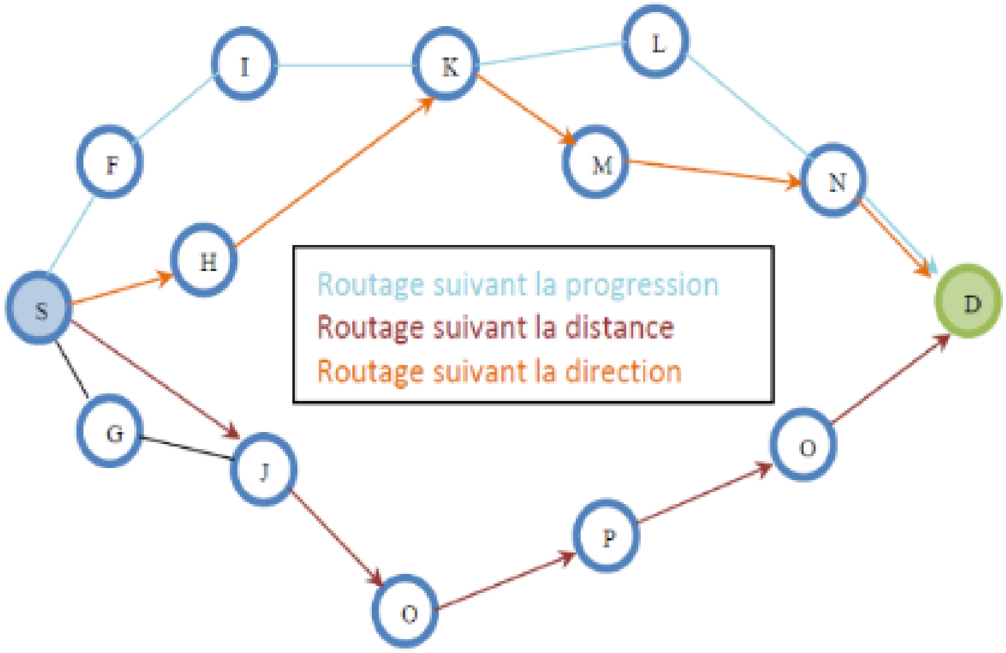


Figure 1.8: Location-based routing [4]

Some Location based routing: [20]

(1) **Sequential assignment routing (SAR).**

(2) **Ad-hoc positioning system (APS).**

(3) **Geographic adaptive fidelity (GAF).**

(4) **Greedy Perimeter Stateless Routing (GPSR)**

(5) **Graph Embedding for Routing (GEM)**

(6) **Location Aided Routing (LAR)**

(7) **Geographic distance routing (GEDIR).**

(8) **Geographic and energy-aware routing (GEAR).**

1.3.6 Definition of The lifetime of a sensor network

The network's operational duration is commonly referred to as its lifetime. In simpler terms, the lifetime of a network represents the period in which it remains functional and capable of executing its intended tasks [21].

1.3.7 Energy consumption in Sensor node

1.3.7.1 Sensing Unit Energy

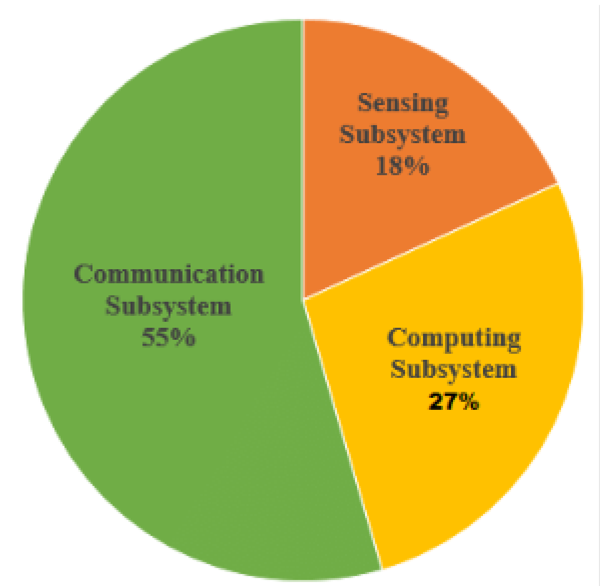
Sensors can include multiple sensors (humidity, heat, motion, position, etc.). Therefore, the sources of energy consumption for a capture operation are sampling, Analog-to-digital conversion, signal processing, and capture probe activation. [22]

1.3.7.2 Processing Unit Energy

The power consumption of the compute unit is slightly higher than the power consumption of the sensor. It is divided into two parts. The first is the switching energy, which is determined by the supply voltage and the total switched capacitance at the software level. The second is lost energy. This is the energy consumed by the computing unit when it is not doing any processing. [23]

1.3.7.3 Communication Unit Energy

This power is the maximum and dominates the power budget at about 60 of the total available energy, mainly used for transmission and reception. The power consumption of this device is highly dependent on the communication protocol. [23]



1.3.8 Factors involved in energy consumption

1.3.8.1 State of the radio module

The radio module is the most energy-intensive component in the sensor node as it ensures communication between nodes. The radio module works in her four modes of operation: idle, transmit, receive, and sleep. - In idle state: The adio is on but not in use. In other words, sensor nodes do not send or receive. - In the transmission state: the radio transmits a packet. - In the receiving state: The radio receives packets. - In the sleep state: the radio is turned off. It has been observed that in most radios, idle mode unnecessarily listens to the transmission channel, resulting in significant power consumption roughly equivalent to that in receive mode. So it makes more sense to turn off the radio completely than to leave it idle when not in use to send or receive data. [24]

1.3.8.2 Access to the medium of transmission

The MAC layer plays an important role in coordinating between nodes and minimizing power consumption. This section analyzes the main causes of power consumption at the MAC layer. [1] Retransmission Sensor nodes typically have a single radio antenna and share the same transmission channel. In addition, transmitting data from multiple sensors at the same time can lead to collisions and loss of transmitted information. Sending lost packets can cause significant energy loss [1] Active listening Active listening of the channel for packets received that may not have been received (idle listening) can lead to a significant loss of node power capacity.

To work around this issue, the node should be in sleep mode for an extended period of time. [1] Overhearing The phenomenon of overhearing occurs when a node receives an unintended packet. Overhearing leads to further loss of energy as the received data includes other sensors. . [1] Overload Several Mac layer protocols work by exchanging messages from the controller (overhead) to provide various functions. Signage, connectivity, access planning, collision. All these messages require additional energy. [1] Over-emission over-emitting occurs when a sensor node sends data to a receiver that is not ready to receive it. In fact, sent messages are considered useless and consume extra energy. [1] Packet size Package size affects power consumption. In other words, reducing the packet size increases the number of control packets exchanged, creating overhead. Otherwise, large packet sizes require the use of high transmit power. [24]

1.3.8.3 Radio Propagation Model

A propagation model represents an estimate of the average power of a radio signal at a specific distance from the transmitter. Radio signal propagation is generally affected by various phenomena. Reflection, diffraction, and scattering by various objects. In general, the power of the received signal is 1/d n. where d is the distance between transmitter and receiver and n is the path issuer loss. [23]

1.3.8.4 Date routing

Routing in WSN is multi-jump routing. Delivery of packets from a particular source to a destination is done through multiple intermediate nodes. A node, therefore, expends energy to transmit its data or relay data from other nodes. In this context, improper routing policy can severely impact network life. [23]

1.3.9 Applications of Wireless sensor networks

The cases of the application are very numerous :

1. **Military Application**
   WSNs are becoming an integral part of military command, control, communications, and intelligence systems. Sensors can be deployed on the battlefield to monitor the presence of troops and vehicles and track their movements, allowing close surveillance of opposing forces. [25]

Figure 1.9: Military Application

2. **Environment Monitoring**
Controlling environmental parameters via sensor networks can lead to several applications. For example, using heat sensors in forests can help identify potential fire outbreaks, making it easier to deal with forest fires before they spread. Placing chemical sensors in urban environments can help detect pollution and analyze air quality. Likewise, industrial use prevents industrial risks such as leakage of toxic products (gases, chemicals, radioactive elements, oils, etc.). [9]

Figure 1.10: Environment Monitoring

3. **Medical Application**
Wireless sensor networks enable constant monitoring of vital human organs with microsensors that can be swallowed or implanted under the skin (monitoring blood sugar, detecting cancer, etc.). It can also facilitate the diagnosis of some diseases by taking physiological measurements such as blood pressure and heart rate, etc. [26]

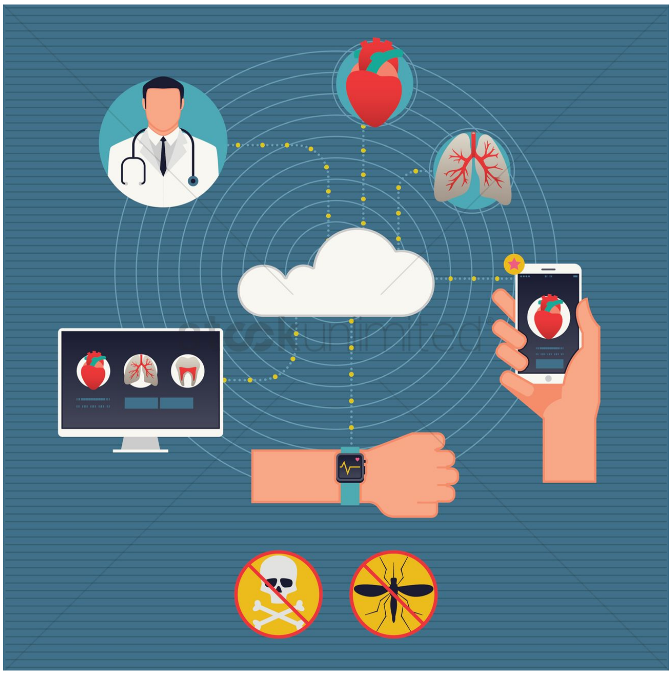


Figure 1.11: Medical Application

4. **Home Intelligence**

   The use of motion and temperature sensors in future so-called smart homes will enable the automation of some domestic operations, such as turning off the lights or turning off the music when the room is empty. Air conditioning and heating are adjusted according to multiple measurement points, and an anti-theft sensor triggers an alarm if an intruder tries to break into the home. [27]

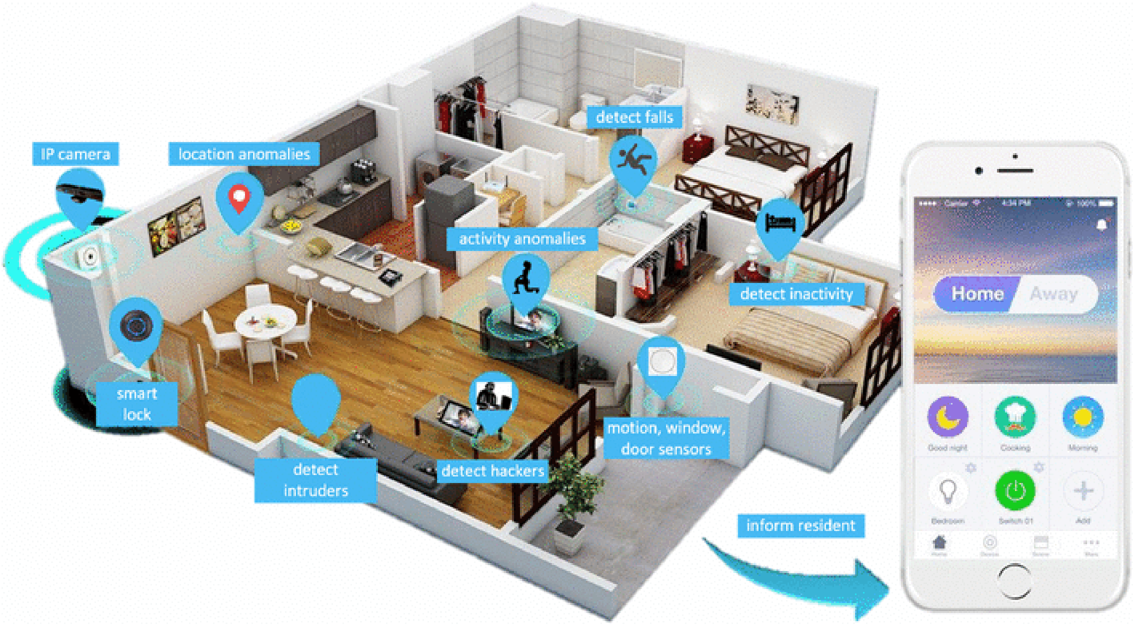


Figure 1.12: Home Intelligence

5. **Agricultural Applications**
The use of wireless sensor networks in the agricultural industry is becoming increasingly common. With wireless networks, farmers no longer need to maintain wires in difficult environments. Pressure transmitters can be used to monitor water tank levels, wireless I/O devices can be used to control pumps, and water usage can be metered and sent wirelessly back to a central control center. [25]

Figure 1.13: Agricultural Applications

1.3.10 Advantages and Disadvantages of WSN

1. **Advantages** [28]

   - It is suitable for hard-to-reach places such as the sea, the mountains, the countryside, and in deep forests.
   - New devices can be added at any time.
   - It also gives flexibility if it happens to need more workspace.
   - It can also be accessed from the central monitor.

2. **Disadvantages** [28]

   - It's less secure because hackers can break into access points and grab all data.
   - Slower and more complex than wired networks.
   - Easily disturbed by the surroundings (walls, microwaves, long distances due to signal attenuation, etc.).
   - Communication speed is relatively slow.
   - Determining Battery Life and Transmission Capabilities [29]

1.4 Conclusion

This chapter introduced the basic concepts of wireless sensor networks, their architecture, and a wide range of applications, as well as general routing protocols and their taxonomy, and we talk also about energy of sensors, finally we terminate by some applications and advantages in real life .

Since energy plays an important role, the next chapter will discuss about Clustering technique help for maximize network's life .

Chapter 2

Clustering in Wireless Sensor Networks (WSN)

2.1 Introduction

Wireless sensor networks are growing in popularity due to major advances in microelectronics that have greatly reduced the cost of manufacturing sensor nodes.

A wireless sensor network (WSNs) is an infrastructure-less network consisting of multiple widely distributed autonomous sensors using powerful base stations (BS) to gather information from these sensor nodes. As mentioned earlier, these sensors suffer from resource scarcity due to power limitations, making balancing performance and maximizing network longevity a major challenge for wireless sensor networks.

Clustering is an important solution for extending the lifetime of sensor networks, it consists of dividing the network into geographically homogeneous regions with the main goal of minimizing energy consumption and consequently increasing network longevity.

In this chapter, we will describe the definition of Clustering, its attributes and its characteristic, benefits, and goals of clustering, and we will present also, clustering techniques and various clustering algorithms in wireless sensor networks.

2.2 Definitions

2.2.1 Clustering

Clustering is a technique for dividing a network into clusters, and each group designates one of these nodes as its leader, called the cluster leader, and the latter as the cluster leader,Communicate with other nodes in the group and cluster leaders in other groups [15].

2.2.2 Cluster

A cluster is a group of sensor nodes that form a network organizational unit. Due to the high density of these networks, it is necessary to divide them into groups to simplify the communication task and to deal with various constraints. [30]

2.2.3 Cluster Head (CH)

A special node called a cluster head (CH) is responsible for organizing and coordinating, aggregating and/or processing the data and communications collected within the cluster. That selection is done in a deterministic or random way thanks to a specific metric or combination of metrics. [31]

2.2.4 Base Station (BS)

A base station is at the top of a hierarchical WSN. Establish communication links between sensor networks and end users. [32]

2.2.5 Member Node

A member node in a cluster is neither the cluster head nor a base station directly connected to the cluster leader. [31]

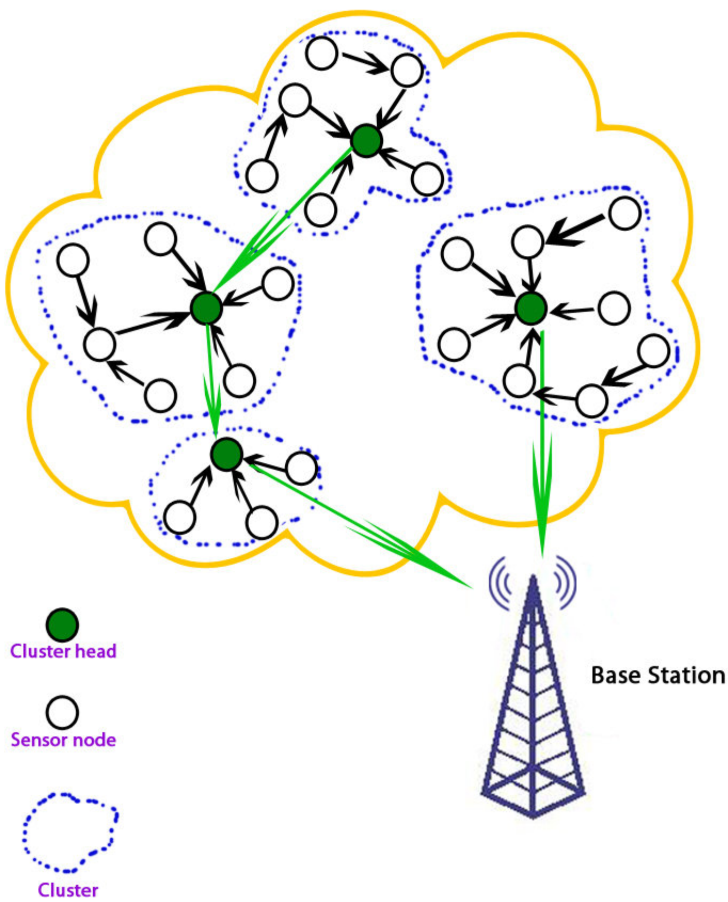


Figure 2.1: Clustering Architecture [5]

2.3 Clustering attribute taxonomies

Different clustering features are used to classify different clustering techniques. the three characteristics of clustering are: [33]
- Cluster properties
- CH properties
- Clustering process properties

2.3.1 Cluster properties

2.3.1.1 Cluster count

In some practical approaches, the set of CH is pre-determined and thus the number of clusters is pre-determined. Random selection of CHs from provisioned nodes typically creates a variable number of clusters. [33]

2.3.1.2 Cluster size

2.3.1.3 Intra-cluster communication

Some clustering approaches are based on direct communication between nodes and their CHs. However, multi-hop connections from nodes to CHs may be required, especially if the communication range is limited or the number of CHs is limited. [24]

2.3.1.4 Inter-cluster communication

Inter–cluster communication (communication between clusters) can be direct or multi-hop communication. Multi-hop mechanisms are usually preferred for energy-efficient data transfer from CHs to BS via intermediate CHs within a large WSN. In some small WSN applications, the communication between CH and BS is single hop transmission. [33]

2.3.2 CH properties

2.3.2.1 Mobility

The mobility of the cluster head node affects member nodes and requires regular maintenance. A fixed cluster head, on the other hand, aims to form a stable cluster and makes the network easier to manage. However, cluster heads may relocate themselves at limited distances to improve network performance. [30]

2.3.2.2 Node Types

We generally refer to the sensor node as CH among the deployed sensors, but sometimes the sensor node with more computing and communication resources is chosen as CH. [34]

2.3.2.3 Role

Some of CH's main tasks are simply forwarding traffic and aggregating or merging the data it collects. [34]

2.3.3 Clustering process properties

2.3.3.1 Method

There are two clustering methods: centralized and distributed. In the centralized approach, a central authority such as a BS controls the entire operation (clustering, CH selection, etc.), whereas the distributed approach has no central authority and is widely used in large WSNs. [33]

2.3.3.2 CH Selection

There are two ways to select CH in WSN: Probabilistic and deterministic methods. In the probabilistic approach, CH is chosen randomly without prior consideration. Deterministic methods select CHs using various metrics such as residual energy, node degree, node centrality, expected residual energy, and distance to BS. [33]

2.3.3.3 Several objectives

Several goals were pursued to form clusters, fault tolerance, load balancing, network connectivity, sociability, and routing [24], or minimization of energy consumption [30]

2.3.3.4 Complexity of the algorithm

The complexity can be constant depending on the number of CHs and sensor nodes [24].

2.4 Clustering characteristic

2.4.1 Type of Clustering algorithm

The clustering algorithm used several algorithms are proposed, there are two types:

1. **C**entralized

   This algorithm runs in a node (usually a base station) that has a global view of the network. This type of algorithm is seldom used because the overhead is generated after the transfer is performed so that the global view of the network and the dynamics of the topology can be maintained, and this view is highly variable.

2. **D**istributed

   The algorithm runs cooperatively on each node in the network. Control tasks are synchronized by exchanging control messages. These types of algorithms minimize the communication associated with storing a global view of the network. This is because each node decides its role independently of others and publishes its decision by sending messages. However, the effectiveness of these algorithms depends on the size and number of these synchronization messages. [35]

2.4.2 The election of the Cluster Heads

The cluster head node consumes more power than other nodes on the network. A cluster head coordinates the actions of its cluster member nodes and aggregates their data. This consumes more power and creates an imbalance in power distribution within the network. To avoid this problem, this cluster head role rotation is organized within the cluster or across the network. Rotation is performed periodically or depending on the power consumption of the cluster head node. [35]

2.4.3 Intra-cluster communication

Communication between the cluster head node and other cluster member nodes can be either single-hop or multi-hop. In direct (one-hop) communication, data packets are sent directly to the cluster head. This assumes that member nodes can reach the CH using transmissions strong enough for good data reception. This type of communication consumes a lot of power when the distance between CH and the node is long. [36]

2.4.4 Inter-cluster communication

A cluster head communicates with a base station either directly or over two hops through a node commonly called a "Gateway node". These nodes can be CH or member nodes of a cluster. Using multi-hop communication reduces energy consumption and increases network scalability. [36]

2.4.5 The level of data aggregation

Depending on the type of sensor used, data aggregation occurs at each node in the network or only at the cluster head level. Data aggregation reduces the size of data exchanged between nodes, resulting in reduced power consumption. Several aggregation techniques are used. That is: Elementary functions such as sum, average, standard deviation, or more complex functions specific to the application used. [35]

2.4.6 types of clusters

The clustering algorithm used can produce two types of clusters: Discrete clusters and interconnected clusters. In the first type, a node can only belong to one cluster at a time (the most common case). Excludes certain applications such as inter-cluster routing, node locations, and synchronization. It uses interconnected clusters. In this type of clustering, nodes can belong to one or more clusters at the same time. [35]

2.5 Benefits of Clustering at WSN

2.5.1 More scalability

In a clustering routing scheme, nodes are divided into different clusters with different levels of assignment. CH is responsible for communication, Member Nodes (MNs) are responsible for event detection and information gathering within the environment. Compared to flat topologies, this type of network topology is easier to manage and more scalable to respond to environmental events. [35]

2.5.2 Fault tolerance

Sensor nodes can be affected by power shortages, transmission errors, hardware malfunctions, malicious attacks, and more. Sensor networks are prone to failure. Fault tolerance is therefore a major challenge for wireless sensor networks. [35]

These types of applications typically require cluster head fault tolerance to avoid severe data loss from critical nodes. Therefore, WSNs need to develop an effective fault-tolerant approach. Grouping into clusters is the most intuitive way to recover from cluster failures, but it usually interrupts ongoing operations. Backup CH allocation is a viable scheme to recover from CH failure [35]

2.5.3 Maximizing network life

Network lifetime is an unavoidable consideration in WSNs, as sensor nodes are limited in terms of power, processing capacity, and transmission bandwidth. The idea of saving energy is to use sensors to effectively cover selected areas and sending the data using clustering algorithms to find ways to extend the life of the network when communicating between groups. [35]

2.5.4 Quality of service

WSN applications make Quality of Service (QoS) requirements. In general, effective sampling, low latency, and transient accuracy are required. It is difficult for routing protocols to meet all QoS requirements because some requirements may violate one or more protocol principles. Existing clustered routing approaches in WSNs are primarily focused on improving power efficiency and not on supporting QoS. Many real-time applications need to consider QoS metrics. [35]

2.6 Clustering Techniques

Clustering methods fall into four main groups: Hierarchy, partitioning, density, and model methods.

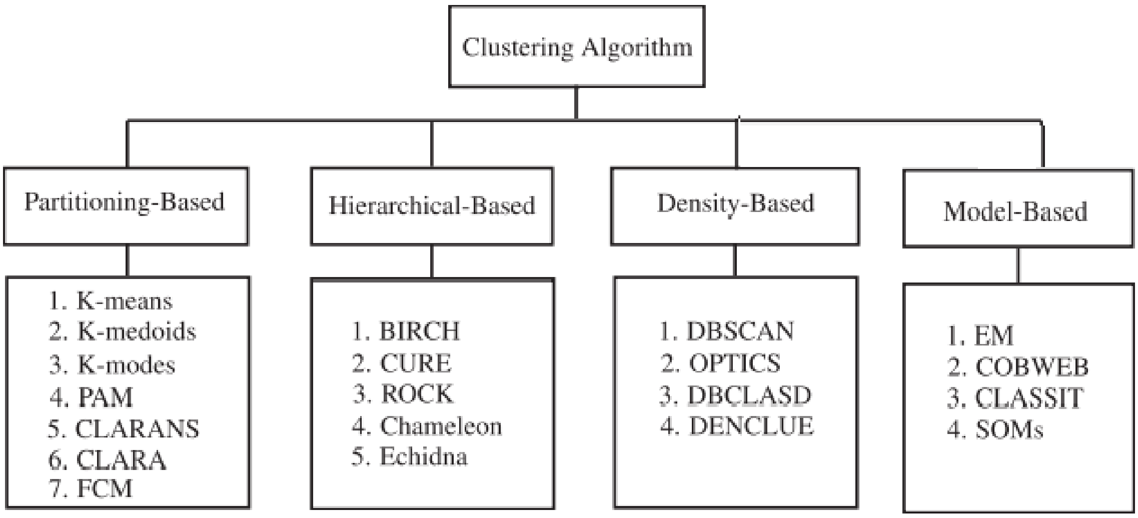


Figure 2.2: Clustering Techniques [6]

2.6.1 Hierarchical-based

Hierarchical clustering algorithms group data objects to form a tree-like structure. It can be broadly divided into agglomerative hierarchical clustering and divisive hierarchical clustering. [37]

The so-called bottom-up agglomeration approach (CHA) merges until we get a single set containing all the objects in the dataset, while the top-down so-called split approach (CHD) takes the dataset and divides it into clusters. which is recursively split. [38]

Examples of these algorithms are LEGCLUST, BRICH (Balance Iterative Reducing and Clustering using Hierarchies), and CURE (Cluster Using Representatives) [37]

2.6.1.1 CURE (Clustering Using REpresentatives)

CURE is an agglomerative hierarchical clustering algorithm. It uses a combination of random sampling and partitioning to handle large databases. First, a random sample drawn from the dataset is split and then each split is partially clustered. Subclusters are reclustered into the desired cluster on the second pass. Experiments confirm that the quality of clusters generated by CURE is far superior to those detected by other existing algorithms. [39]

2.6.2 Partitioning-based

Partitional clustering is very different from hierarchical approaches that generate incremental levels of clusters with iterative merging or splitting. Partitioning clustering assigns a set K of objects into unhierarchical clusters . [40]

Partitioning algorithms are a good choice when dealing with large data sets. However, these algorithms are relatively computationally intensive in terms of clustering consistency, making this approach less effective than the agglomeration approach. [40]

There are many methods of partitioning clustering such are K-means Methods, PAM(Partitioning Around Medoids), CLARA (Clustering LARge Applications) [41]

2.6.2.1 K-MEANS

The principle of this algorithm is that each group's centroid is recalculated for each new individual introduced into the group, rather than waiting for all individuals to be assigned. It is a vector quantization algorithm that attempts to separate a set of points into K clusters. K-Means, defined by McQueen, is one of the simplest algorithms for automatic data classification. The main idea is to randomly choose a set of a priori fixed centers and iteratively search for the best partition. Each individual is matched to the closest center based on Euclidean distance. After associating all data, the mean for each group is calculated to represent the new representative of the group. Once in a steady state (no data changes in the group), the algorithm stops. [38]

2.6.3 Density-based

Density-based algorithms can detect clusters of any arbitrary shape, thus ensuring noise isolation and preventing the formation of irrelevant clusters. These algorithms group objects according to certain density functions. Density is usually defined as the number of objects in a given neighborhood of a data item. With this approach, the size of a given cluster keeps growing as long as the number of nearby objects is above a certain threshold. [42]

Density-based clustering techniques mainly include three techniques: DBSCAN (spatial clustering based on application density with noise), OPTICS (control points to identify the clustering structure), DENCLUE (DENsity CLUstering). [38]

2.6.3.1 DBSCAN (spatial clustering based on application density with noise)

It was proposed by Martin Ester, Hans-PeterKrigel, JorgSander, and Xiawoei in 1996. One of the most popular density-based algorithms. If an object belongs to a cluster, the neighborhood density of the object should be high enough. A cluster skeleton is created using this set of core objects, with overlapping neighbors. The points within the neighborhood of the core object represent the cluster boundaries, the rest are just noise. It requires two parameters [40]:

1. Eps is the starting point.
2. Minpts is the minimum number of points required to form a dense region.

2.6.4 Model-based

Model-based clustering methods aim to enhance the correspondence between data and a mathematical model. These approaches typically assume that the data originate from a combination of latent probability distributions. There are two primary approaches in model-based clustering methods: statistical (probabilistic) and neural network-based approaches [38].

In the probabilistic category, an initial cluster head or other random selection procedure is determined based on the a priori probability assigned to each sensor node. The probabilities initially assigned to each node play a crucial role in individually determining their eligibility as cluster heads [43].

Popular Probabilistic Clustering: Expectation-Maximization (EM), Gaussian Mixture Model (GMM)).

2.7 Related works

2.7.1 Low-Energy Adaptive Clustering Hierarchy(LEACH)

LEACH is one of the first clustering routing approaches in WSNs proposed by Heinzelman et al [44], LEACH is a fully distributed approach and does not require global network information. [45] In LEACH, nodes are organized into a local cluster, with one node acting as the cluster head and the others acting as member nodes. All member nodes send data to their respective CHs, and when data is received from all member nodes, the cluster head performs signal processing functions on the data and sends the data to the remote BS. [46] The main goal of LEACH is to balance the energy consumption of all nodes. Cluster heads are selected by rotation. In this way the power load of the cluster head is evenly distributed to all nodes and the high power consumption of connecting to the base station is distributed and shared to all sensors at a close rate. [?]

The operation of LEACH is divided into rounds, and each round is divided into two phases, a set-up phase and a steady-state phase. The setup phase organizes the cluster, and the steady-state phase distributes the data to the BSs. [45]

The cluster head changes randomly over time to balance power consumption across all nodes. This decision is made using a random number between 0 and 1. A node becomes the cluster head for the current round if the number is less than the following threshold: [44]

$$T(n) = \begin{cases} \dfrac{P}{1 - P\,(r \bmod (1/P))}, & \text{if } n \in G, \\ 0, & \text{otherwise,} \end{cases}$$

where

P is the desired percentage of cluster heads,

r is the current round,

G is the set of nodes that have not been elected cluster heads in the last 1/ P rounds. [44]

After CHs are chosen, a decision is made by each normal node to determine the appropriate CH to connect to form a cluster. CH sends advertisement messages to all other nodes. [47]

In the second phase (steady state), data is sent from the member nodes to the cluster head and then to the base station. [48]

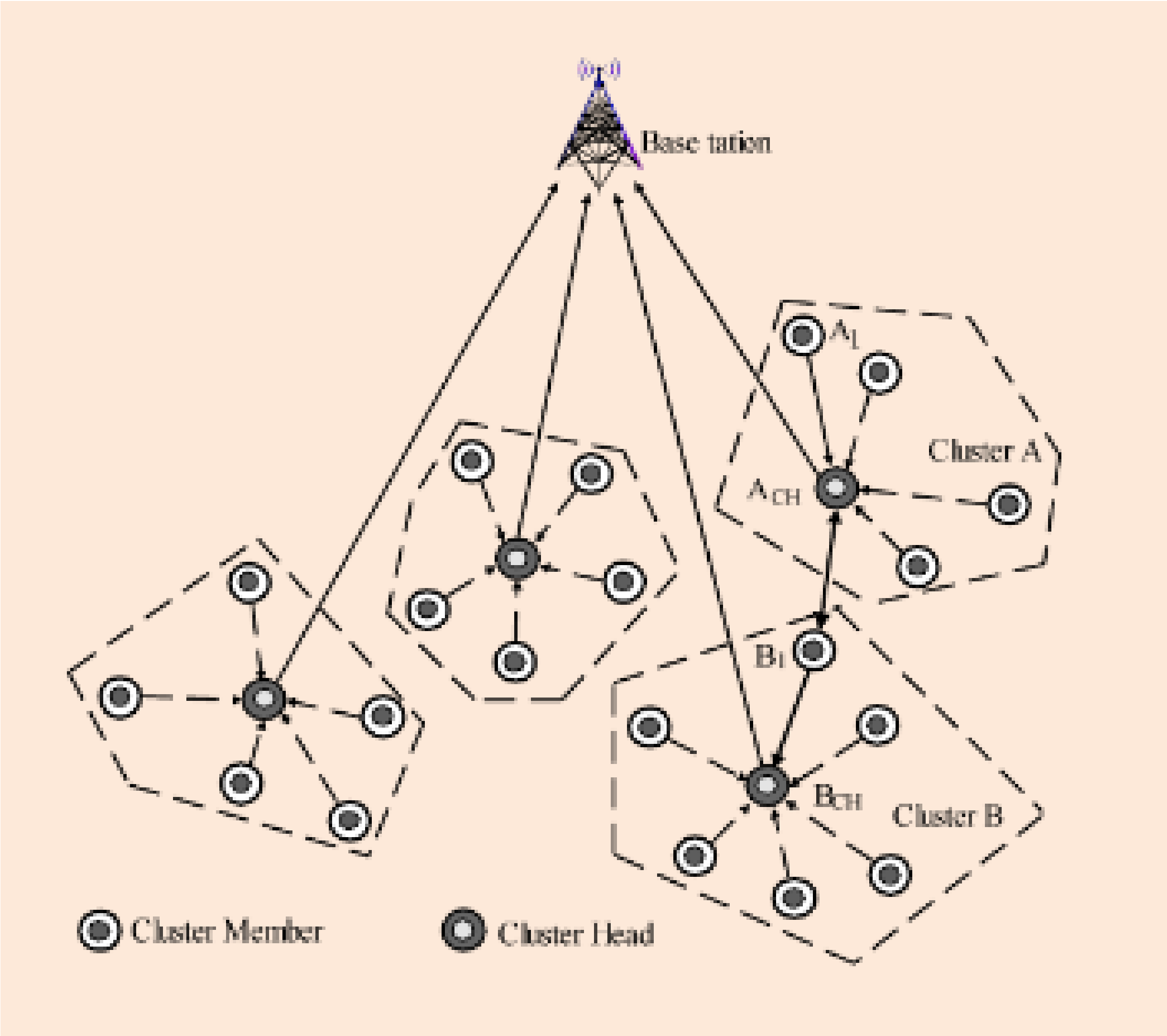


Figure 2.3: LEACH Clustering Structure

2.7.1.1 Advantages [49]

1. Reduce control messages overhead.

2. Low complexity algorithm.

2.7.1.2 Disadvantages [49]

1. Non-uniform distribution of cluster heads.

2. Select the cluster head without considering the remaining energy.

2.7.2 Hybrid Energy-Efficient Distributed Clustering(HEED)

Hybrid Energy-Efficient Distributed clustering (HEED), was introduced by Younis and Fahmy.

Hybrid, Energy-Efficient, Distributed (HEED) is an equal-size clustering protocol that creates equal-sized clusters. The selection of CH in HEED is based on the residual energy of the sensor node and one of the following parameters: Node degree or distance from an adjacent node to CH. Cluster formation in HEED occurs in three stages: initialization; iteration and finalization. During the initialization phase, each node is assigned a probability of becoming a temporary cluster head. This is done using the following formula: [50]

$$CH_{prob} = C_{prob} * \frac{E_{residual}}{E_{max}}$$

where Cprob is the initial probability (i.e., a predefined value)

Eresidual is the residual energy

Emax is the maximum energy of the sensor nodes.

During the iterative phase of HEED, some nodes are closed Provisional cluster headers. If the node is within the communication range of several temporary cluster heads, choose the one with the lowest cost. If a node is not within range of the interim cluster head, it will eventually become the cluster head.

In the finalization phase, there will be one sensor node that has not elected a cluster head. [50]

HEED creates balanced clusters of equal size. size Cluster distance is independent of distance from BS. Due to inter-cluster communication (that is, forwarding traffic between CHs), sensor nodes near the BS quickly consume power. In fact, while routing normal communication within traffic (i.e. intra-cluster traffic), it is also burdened with heavier route traffic from the rest of the network nodes. This relay traffic not only impacts the network lifetime but also leads to network partitions (near base stations). Railway station). This is also called a hot spot problem. To address this issue, unequal clustering was developed. [50]

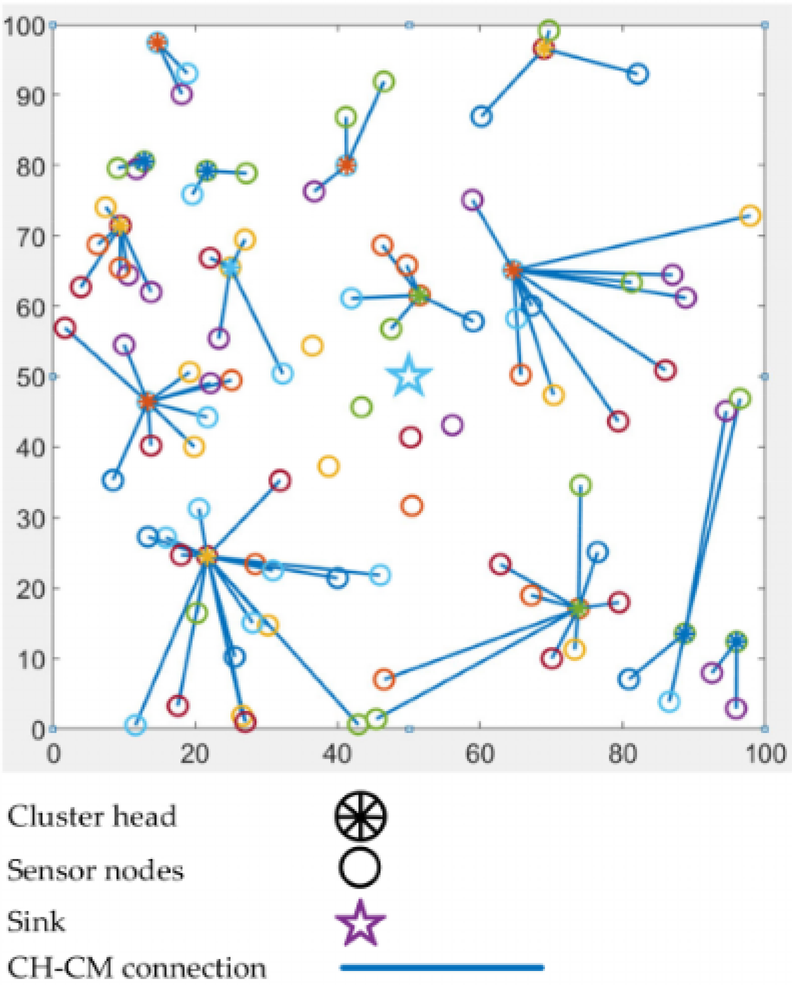


Figure 2.4: HEED Clustering Structure [7]

2.7.2.1 Advantages [51]

1. Applicable to large-scope networks.
2. The clustering process can be completed in a fixed number of iterations, with cluster heads fairly evenly distributed across the network.

2.7.2.2 Disadvantages [51]

1. Probabilistic cluster head selection methods do not allow for a real even distribution of cluster heads in the network.
2. Multiple iterations are required to form a cluster, and each iteration broadcasts many packets, resulting in overhead.

3. Some cluster heads, especially near sinks, can experience premature outages, creating hot spot in the network.

2.7.3 Power-Efficient Gathering in Sensor Information Systems (PEGASIS)

Proposed by Lindsey et Raghavendra, PEGASIS is an improved version of the LEACH protocol. [15]

Each sensor node in this network is randomly placed and uses a greedy algorithm to draw a long chain of all sensor nodes. This process is called "chain Construction". Also, the "leader" node chosen during the "chain Construction phase" could be the node closest to the BS or sink. Token passing can also be used to implement data aggregation. Each sensor node in the chain forwards its data to neighboring nodes, and the next node merges its data with the previous node and sends the data to the next node. Eventually after merging, the data reaches the leader node, which combines the data and brings it to the sink (base station). The token is sent in one direction from the "leader" node to the end of the chain, and the node sends its data to the nearest node in the same direction as the leader node. [52]

The next node merges the received data with its own data and sends it further towards the master node. Systematically, data is received at the reader node, and the "reader" node compresses the data and sends it to the sink. This distributes network energy evenly, so you don't starve nodes when you don't need them. Only the leader node "Aggregating" data and sends it to the sink, so other nodes don't have to waste energy on non-leader "data aggregation". [52]

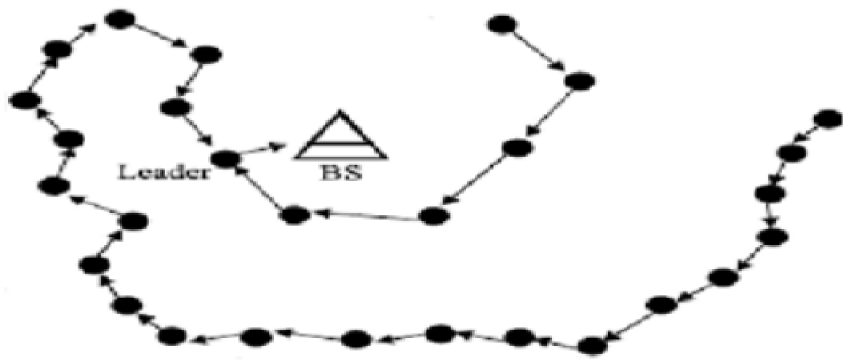


Figure 2.5: PEGASIS Clustering Structure

2.7.3.1 Advantages [49]

1. Uniform distribution of energy conception
2. Reduce overhead by dynamic clustering

2.7.3.2 Disadvantages [49]

1. Long delay in sending information
2. Week scalability.
3. General networking knowledge required
4. Not suitable for variable network topologies

2.7.4 Chain-cluster based mixed routing (CCM)

CCM takes advantage of the low power consumption of the PEGASIS protocol and the short transmission delay of the LEACH protocol. The CCM protocol has two phases:

1. Phase 1 Phase 1 is known as chain-based routing. In this step, sensor nodes build a chain for intra-connectivity and all chain member nodes send data to the chain head using a chain-based routing concept. This process involves her two steps: Selection a chain head node and transferring data within the chain.
2. Phase 2 In Phase 2, CCM has two steps. cluster header voting and intra-cluster data transfer. All chain heads establish cluster-based routing as connections. Finally, the selected cluster head sends the merged data to the base station. More energy is used to transmit data when the cluster head is further away from the base station. [53]

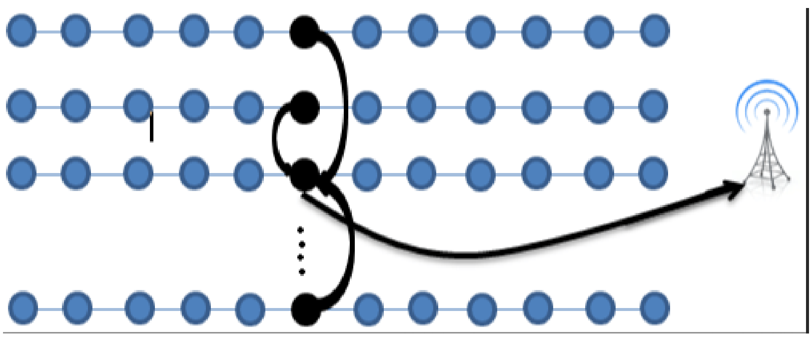


Figure 2.6: CCM Clustering Structure

2.7.4.1 Advantages [49]

1. Reduces the high overhead of cluster head selection compared to LEACH.
2. Reduces long transmission delay compared to PEGASIS.

2.7.4.2 Disadvantages [49]

1. If CDM is not supported, the transmission delay will increase.
2. Unsuitable for large-scale networks.

2.7.5 Hybrid Clustering Tree-based Energy(HCTE)

This algorithm uses two cluster heads to balance the energy of each cluster. Each of these cluster leaders in each cluster is responsible for different tasks. The algorithm also uses a multi-hop transmission mechanism to route data from the cluster head to the base station. [49]

HCTE, like other cluster-based routing algorithms, has multiple phases for structuring the network in clusters and sending data to sinks. For this reason, HCTE has five phases: Initial cluster head announcement, cluster formation, second cluster head announcement, schedule creation, and data transmission. [54]

1. Initial Cluster Head Announcement the initial cluster head announcement phase, is almost identical to HEED's cluster head selection algorithm. [54]
2. Cluster Formation

   At this stage, cluster formation is complete and each node tries to find the best cluster head to join. For this purpose, each node calculates a trust level based on perch transfer within the cluster head, competes with other nodes and joins the cluster according to the highest trust level Head. [49]
3. Second Cluster Head Announcement

   At this stage, a second cluster leader was announced. All nodes in the cluster compete for self-trust scores, and eventually the node with the highest trust score introduces itself as the head of her second cluster. (Cluster headers are used for intercommunication). [49]
4. Schedule Creation

   The scheduling phase in HCTE is the same as in LEACH. [54]

5. Data Transmission using multi-hop data transfer. [54]

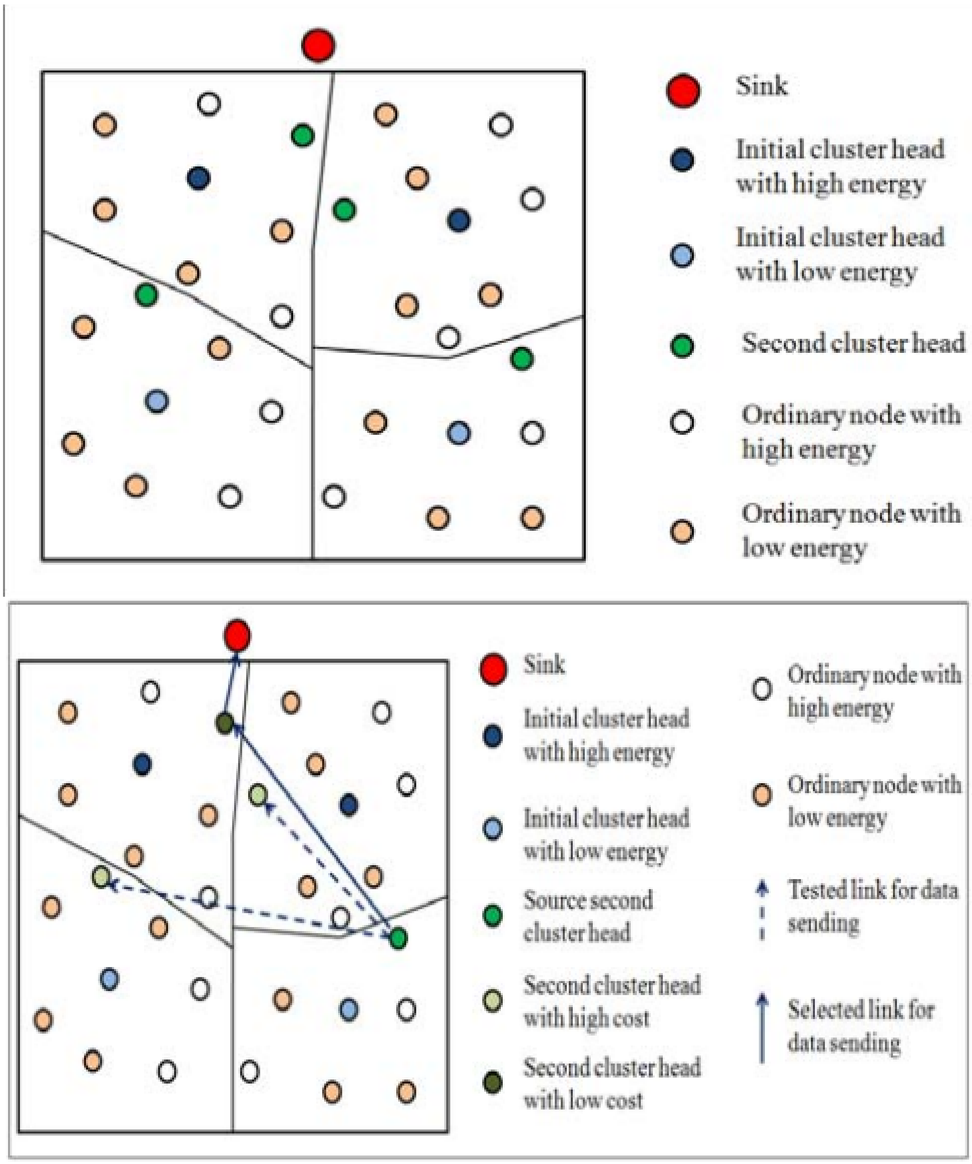


Figure 2.7: HCTE Clustering Structure

2.7.5.1 Advantages

1. Uniform distribution of energy

2.7.5.2 Disadvantages

1. Low scalability

2.7.6 Cluster Head Election mechanism using Fuzzy Logic (CHEF)

The CHEF protocol (Fuzzy Logic Cluster Head Election Mechanism for Wireless Sensor Networks) uses a fuzzy logic approach to maximize the lifetime of WSNs. BS does not need to collect information from all nodes. Instead, the CHEF protocol uses a localized CH selection mechanism using fuzzy logic. Each node chooses a random number between 0 and 1. If this random number is less than Popt, the node uses FIS to calculate chances and announce candidate messages containing chances. This message indicates that the node is a candidate for CH with the value Random. [55] Popt is calculated as:

$$P_{opt} = p \times \alpha$$

Where: p is as in Equation (2.12) and is a constant value that defines the ratio of the candidate for cluster head. The node then listens for candidate messages from nodes within radius r in the equation.

$$r = \sqrt{\frac{area}{\pi . N . p}}$$

The node with the highest probability is chosen as CH. After CH selection, everything (advertising message, CH join message, and steady-state phase) is the same as LEACH. FIS uses two variables. Residual energy and distance between nodes exit, and 9 IF-THEN rules. [55]

2.7.6.1 Advantages [56]

1. Keeping in view the entire network for the CH election
2. Optimal CH selection

2.7.6.2 Disadvantages [56]

1. Excessive message and communicational overhead
2. Expensive CH election process in terms of energy consumption

2.7.7 Weight-Based Clustering Protocols (WCR)

The proposed protocol (WCR) is a cluster-based low-power protocol for wireless sensor networks. The purpose of this protocol is to reduce the power consumption of nodes for routing data to base stations and extend the life of the network. [57] This type is used as a criterion for selecting cluster heads. This weight is used to measure some factor, such as the node's energy or distance from the cluster head, or the number of times a node becomes a cluster head, depending on the algorithm. Each node computes its weight at each cluster iteration. Clusters are created to minimize the power consumption of wireless sensor networks. [58]

2.7.7.1 Advantages [49]

1. Cluster size limitation
2. Node tries to re-connect to the cluster head only if disconnected

2.7.7.2 Disadvantages [49]

1. The overhead of data collection and rapid death of the cluster head node.
2. Excessive amount of computing and communications, and energy consumption.
3. Each node must save all data nodes before initializing the network

2.7.8 Unequal Clustering Size (UCS)

For network configuration, unequal cluster size (UCS) is proposed to balance the power consumption of cluster heads in multi-hop sensor networks and extend the network lifetime. Cluster heads are deployed deterministically and are considered much more expensive (supernodes) than simple sensor nodes. You can move around to adjust locations while managing the size of your cluster and the expected load from other, more distant clusters. In UCS, the detection field is assumed to be circular and divided into two concentric circles called layers. Through theoretical and experimental analysis, the size of clusters in the inner layer should be reduced to achieve more uniform energy consumption. [59]

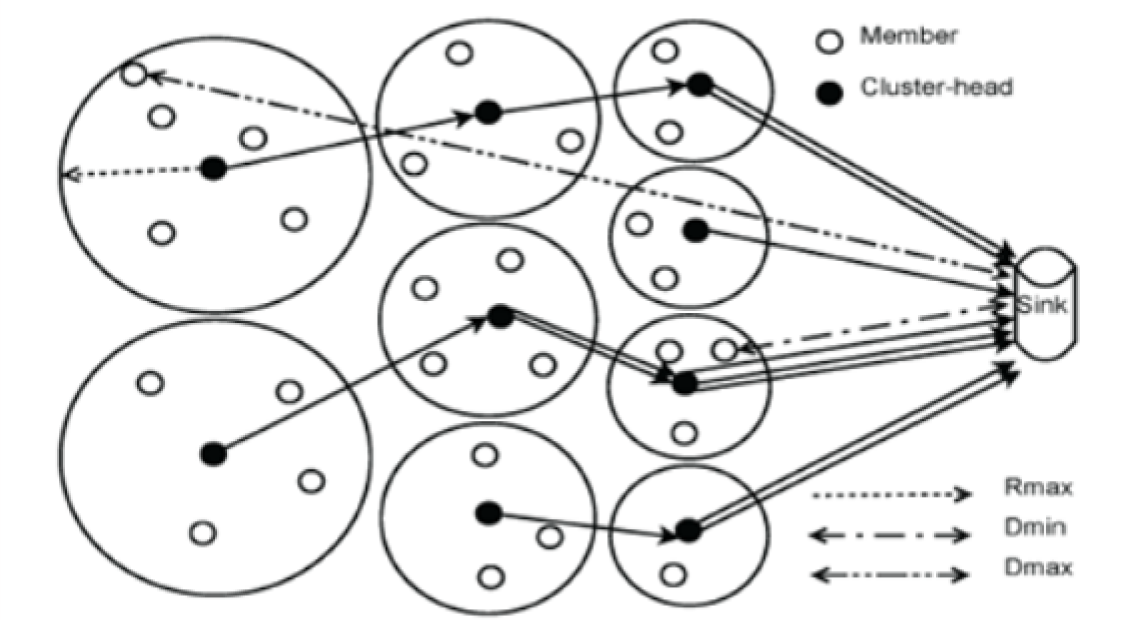


Figure 2.8: UCS Clustering Structure

2.7.8.1 Advantages

1. Flexibility in Cluster Size: UCS allows for the formation of clusters with varying sizes.

2.7.8.2 Disadvantage

the difficulty of use in real-time applications due to predefined assumptions. [60]

2.7.9 Energy Efficient Unequal Clustering (EEUC)

Energy-efficient unequal clustering (EEUC) is proposed by Li et al [61], In order to save more power in intra-cluster and inter-cluster communication, it is proposed to balance the power consumption between clusters where the cluster size near the sink node is much smaller than clusters far from the sink node. I'm here. Here we extend the life of the network and balance the load between nodes. Protocol that is a distance-based scheme, therefore, requires each node to have global knowledge of the sink node. It solves the hot spot problem by arranging cluster size proportional to the distance from the base station. [62]

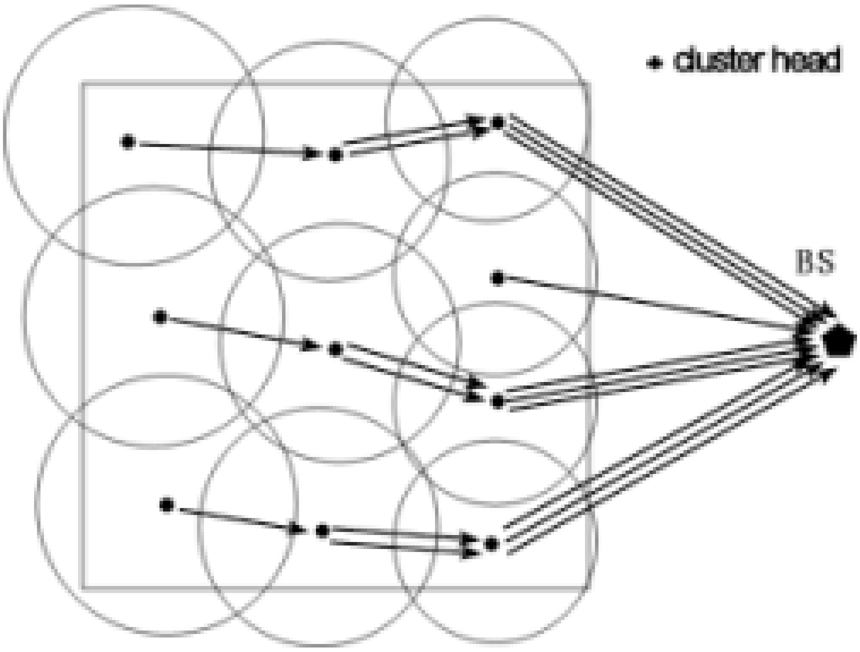


Figure 2.9: EEUC Clustering Structure
[61]

2.7.9.1 Advantages

1. The distance between nodes and base stations was reduced and the network was divided into unequal size clusters. [?]

2.7.9.2 Disadvantage

Select CH for aggregation and simultaneous transmission of data transmission to the base station. [?]

2.7.10 Threshold-sensitive energy efficient sensor network(TEEN)

It is based on the LEACH algorithm. The clustering process uses two thresholds called the soft threshold and the hard threshold. The goal of this threshold is to reduce the amount of data transferred between nodes. A hard threshold is one of the transmitting rules. Data is sent if the value received from the sensor is greater than this threshold. Otherwise, no information will be sent to the base station. Soft thresholds are thresholds that make the algorithm more flexible. If the node's value is less than the hard threshold, but the difference between the node's two most recent values are greater than the soft threshold, the data is sent to the base station. [49]

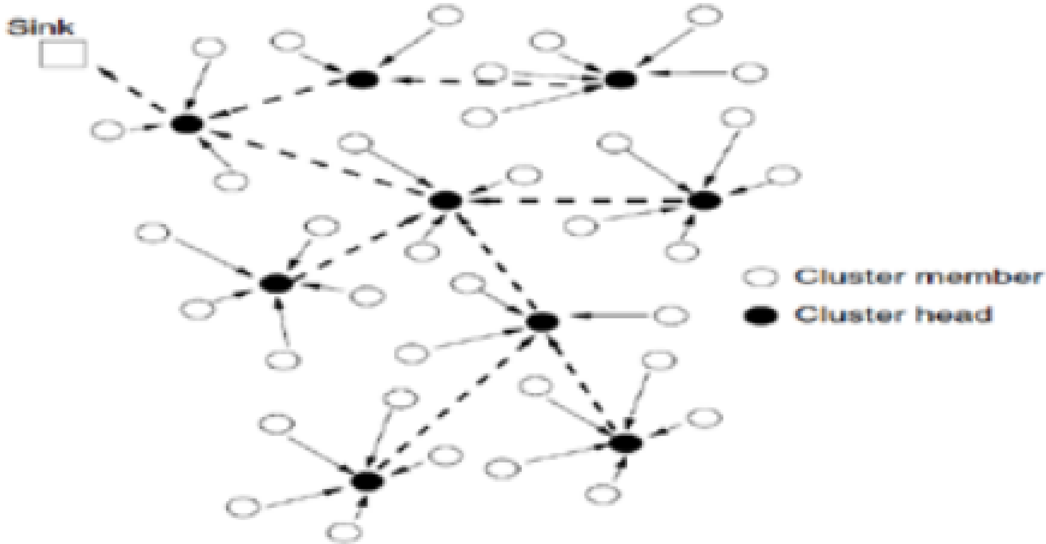


Figure 2.10: TEEN Clustering Structure

2.7.10.1 Advantages

1. The soft threshold is modifiable, allowing the user to change the fresh parameter on each cluster switch if desired. [51]

2. Time-sensitive data reaches users almost instantly. This scheme is therefore ideal for time-sensitive data collection applications. [63]

2.7.10.2 Disadvantages [51]

1. TEEN is not suitable for applications with periodic reports, as users may not be able to retrieve any data if the value of the attribute does not reach the threshold.

2. There is the fact that time slots are wasted and the sink may not be able to distinguish between dead and live nodes.

2.8 Protocols Comparison Table

The comparison between the studied protocols is summarized in the table provided:

Protocol	Classification	Energy Efficient	Scalability	advantages	disadvantages
LEACH	Hierarchical	high	good	Low complexity algorithm	Select the cluster head without considering the remaining energy
TEEN	Hierarchical	high	low	Time-sensitive data reaches users almost instantly	TEEN is not suitable for applications with periodic reports
PEGASIS	Hierarchical	Max	low	Uniform distribution of energy conception	Not suitable for variable network topologies
HEED	Hierarchical	high	Good	Applicable to large-scope networks	Probabilistic cluster head selection methods do not allow for a real even distribution of cluster heads in the network
CCM	Hierarchical	moderate	moderate	reduces long transmission delay	Unsuitable for large-scale networks
HCTE	Hierarchical	moderate	low	Uniform distribution of energy	Low scalability

Protocol	Classification	Energy Efficient	Scalability	advantages	disadvantages
CHEF	Hierarchical	high	moderate	Optimal CH selection	Expensive CH election process in terms of energy consumption
WCR	Hierarchical	high	moderate	Cluster size limitation	excessive amount of computing and communications, and energy consumption
UCS	Hierarchical	high	high	Flexibility in Cluster Size	he difficulty of use in real-time applications
EEUS	Hierarchical	high	moderate	the distance between nodes and base stations was reduced	Select CH for aggregation

2.9 Conclusion

This chapter has provided a overview of clustering in WSN, In this chapter, we have explored the concept of clustering in WSN and its characteristic.

we highlighted the benefits of clustering in WSN ,where the important one is maximize network life.

We delved into various clustering algorithms used in WSN, including LEACH, HEED, and others. We examined their principles, advantages, and limitations.

Chapter 3

Conception

3.1 Introduction

Clustering in WSN is an effective way to structure the network to facilitate communication and data management. This automatically leads to a gin in terms of energy consumed.

This chapter describes the approach we have proposed for clustering in wireless sensor networks(WSN).

The main objective of this work is to evaluate two important clustering techniques in RCSF and develop a new clustering technique that is efficient in terms of the energy and service life of this network.

Our proposal is a combination between LEACH clustering and K-means clustering.

LEACH has many advantages in terms of energy savings and communication efficiency in WSN. However, because LEACH is designed for small to medium-sized WSNs, it may not be as effective in large sensor networks.

Random selection of cluster heads can lead to inequities in workload and power usage, which can affect network performance at scale.

In this context, we try to eliminate the disadvantages of LEACH clustering by combining it with K-means clustering.

K-Means is a simple and efficacies unsupervised method of clustering groups data points into k clusters in which each observation belongs to the cluster with the nearest mean. In this chapter we will present the different components used to implement it. He mainly studies architecture and detailed descriptions of the different entities, as well as their roles.

3.2 Description of the techniques studied

3.2.1 K-MEANS algorithm

Based on geographical location, this phase is responsible for the distribution of the network. This partitioning is carried out according to the distance between the sensor nodes (the nearest nodes are grouped together). Our strategy allows the separation of the network into k clusters by the use of the "k-means" algorithm. It is a simple and effective method. It can be used with large databases (thousands of sensors).

The k-means algorithm aims to compartmentalize the n sensor nodes if $I \in [1, 2, ..., n]$ for k clusters $Pj \in [1, 2, ..., K]$(k<n) using k centers of "Cj" chosen arbitrarily. This algorithm aims to minimize the distance between the "Si" sensor nodes inside each "Pj" cluster using the formula (1):

$$arg\,min_P \sum_{j=1}^{k} \sum_{i=1}^{n} \left\| S_i - C_j \right\|^2 \quad \textbf{(1)}$$

Where ‖ Si- Cj ‖: the distance between the sensor nodes "Si" and the cluster center "Cj".

The method calculates the distance between all Si sensor nodes and all initial "Cj" centers and assigns each node to the nearest cluster. Once all sensors are assigned, new cluster centers are calculated (thus, centers are mobile), where we determine the mean "Gj" (see formula (2))

$$G_j = \frac{1}{|P_j|} \sum_{i=1}^{n} S_i \quad \textbf{(2)}$$

of each cluster and designate this average as a new center "Cj". This process is repeated until partitioning reaches a certain stability (indeed, a threshold can be set as a stopping criterion).

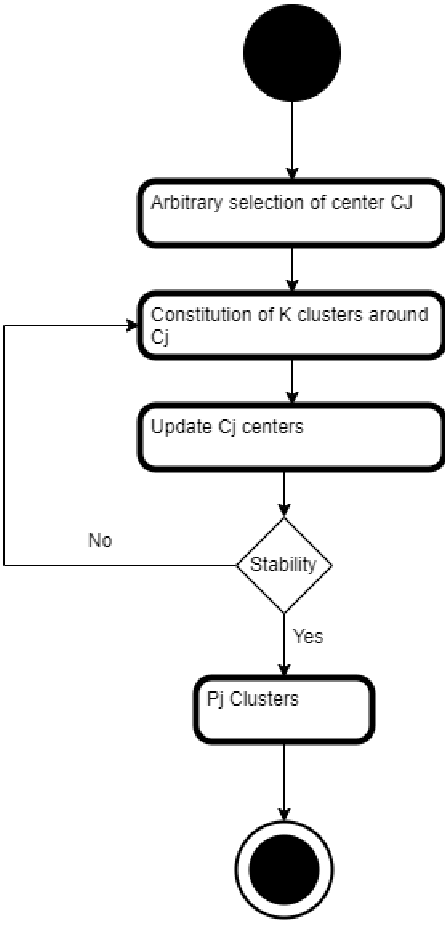
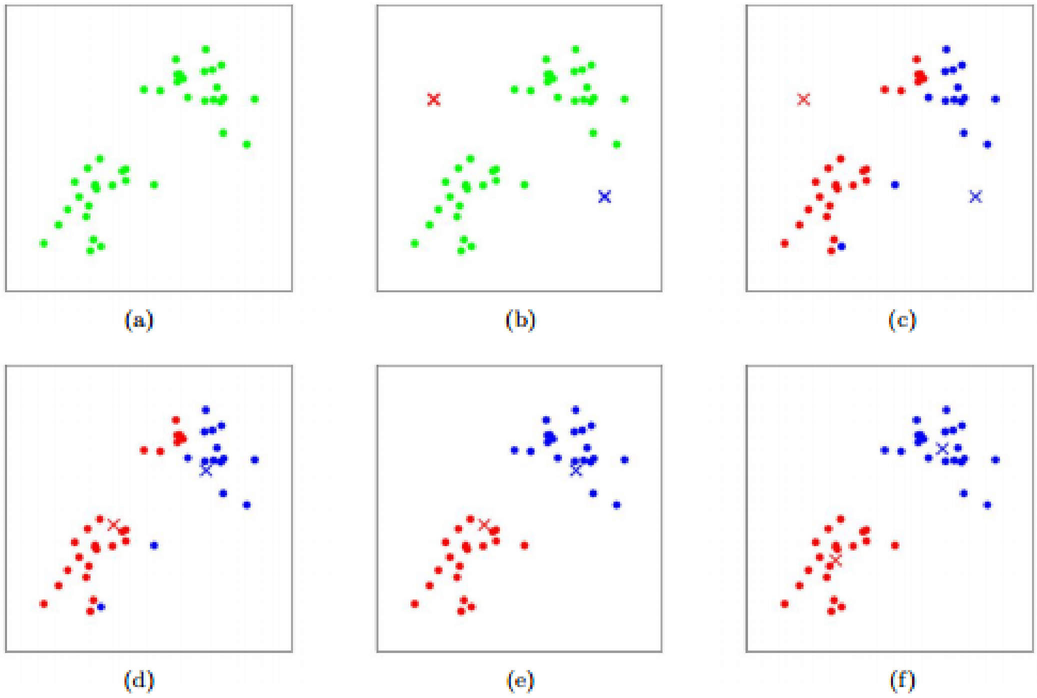


Figure 3.1: Network partitioning process by kmeans method



The advantages of K-means [38]

1. **Flexible:** The K-means algorithm is capable of adapting to various data changes. If any issues arise, adjusting the cluster boundaries allows for quick modifications to the algorithm.

2. **Effective:** This algorithm is particularly effective when it comes to partitioning large datasets. Its success largely depends on the cluster shapes, with K-means performing well in hyper-spherical clusters.

3. **Low compute cost:** In comparison to other classification methods, the k-means technique is known for its fast and efficient computation.

The disadvantages of K-means [38]

1. **Specify the Number of Clusters:** In order for K-means classification to yield accurate results, it is crucial to specify the desired number of clusters (K) at the outset of the algorithm.

2. **Calculation Limitations:** When dealing with large datasets, the application of dendrogram techniques can lead to computer crashes due to the high computational load and dynamic memory constraints.

3. **Inherent Inconsistency:** K-means clustering can produce variable outcomes across different runs of the algorithm. The random initialization of cluster centroids can lead to different results, introducing an inherent level of inconsistency.

3.2.2 LEACH Clustering Standard

The LEACH protocol is used as a probability formula for randomly selecting nodes. The tasks of collecting data and transmitting the data to the cluster heads are performed by ordinary nodes. On the other hand, cluster heads undertake the tasks of processing, aggregating, receiving and transmitting data to sinks or base stations. So cluster heads consume more energy than non-cluster heads. Each node is elected as a cluster head to continue the data transmission process. Energy consumption of nodes proceeds systematically until the last node loses energy and dies. This increases the lifetime of the network. The LEACH protocol is divided into several rounds. Each round consists of the following two phases: [8]

1. **Set-up** In this phase, a random number R, from 0 to 1 is selected randomly from each node with a threshold T value of the node. The threshold can be calculated from the following [8]

$$T(n) = \frac{p}{1-p\left(r\left(mod\frac{1}{p}\right)\right)}$$

   Here, T(n) is the threshold value, P is the optimal cluster heads selection probability, and r is the present round. If the value of the sensor node's random number (R) is more than T, then the given node will not be appointed as CH for the current round. If R is less than T, the node will be appointed as the CH for the current round. After the cluster head selection is completed, the same round of messages will be propagated to other nodes, and these nodes will become cluster heads. Various cluster groups are formed in the network so that the cluster head receives a node request message with a high-power signal, which means that this CH is closer to the node than other CHs. Thus the distance between nodes and cluster heads is minimized and the energy loss in the network is also reduced. [8]

2. **Steady-state** In this phase, CDMA and TDMA techniques are used, so the duration of this phase is longer than the previous ones. In order to avoid conflicts

between signals, each node has a certain time to transmit data, which is allocated by the cluster head. When packets are received by cluster heads, they are assembled without confusion using CDMA and forwarded to sinks. This completes the first round, and the process is repeated until the energy of the remaining nodes is exhausted. Figure 2 illustrates the flow chart of LEACH protocol, beginning with the selection of the cluster heads and ending with sending data to the base station. [8]

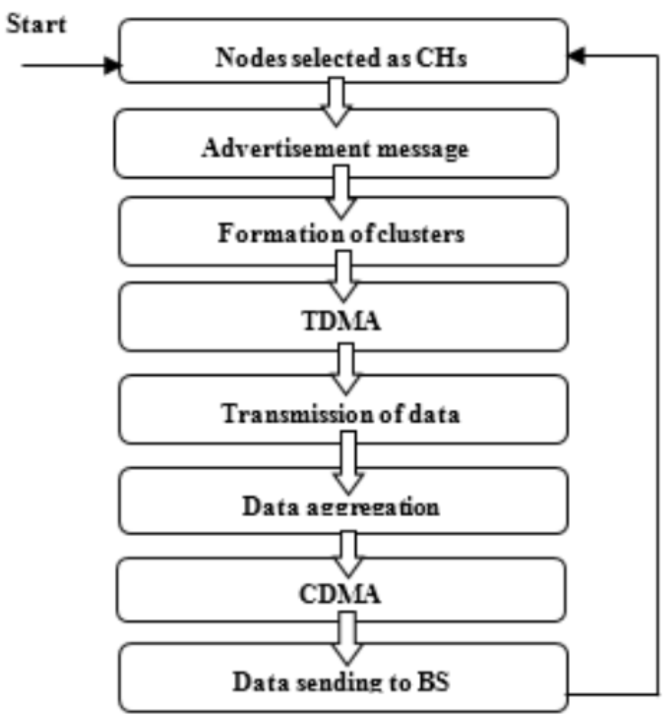


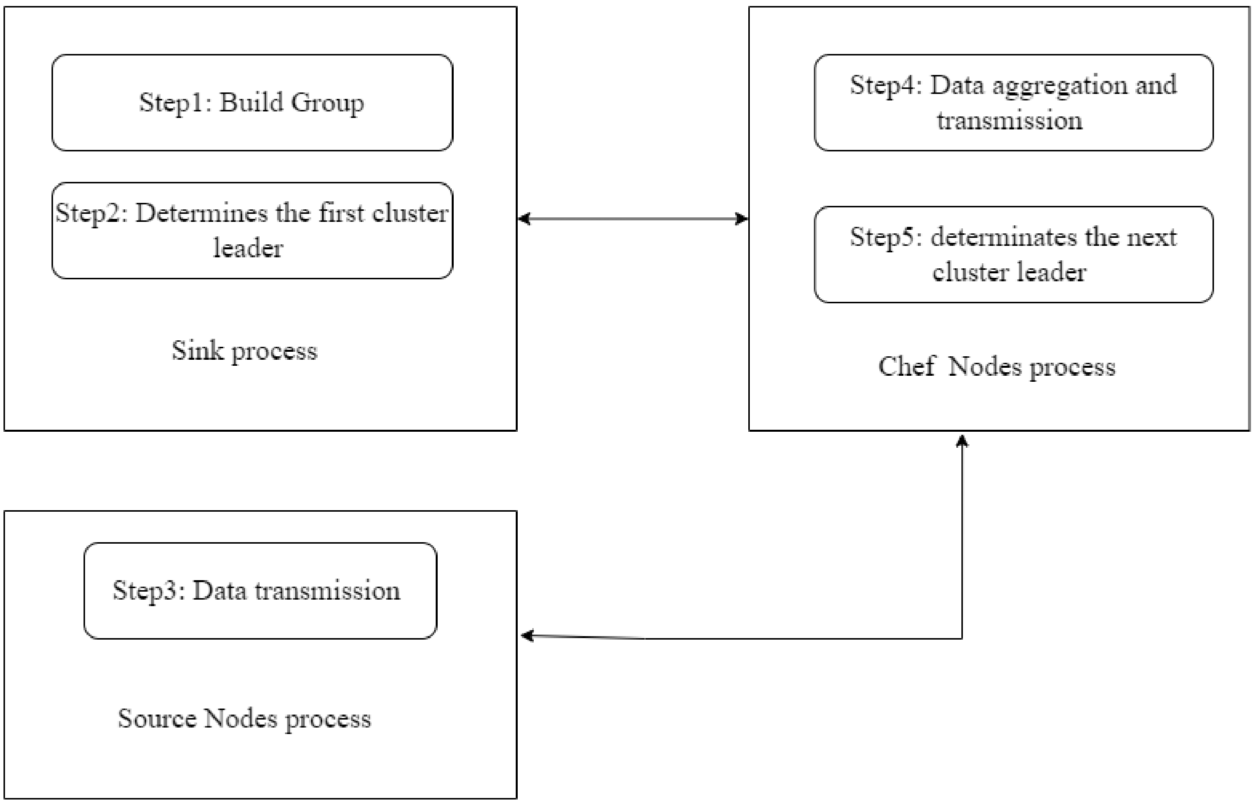
Figure 3.2: LEACH Clustering Steps [8]

3.3 Description of the proposed approach kmeans-LEACH

3.3.1 General description

In order to achieve our goal, we propose a new method (kmeans LEACH) of clustering that encompasses the strengths of LEACH clustering technique and Kmeans clustering technique. Our proposed approach can be classified as an improvement of the LEACH routing protocol by the advancement of the clustering technique for this protocol.

The overall structure of our model is represented by the figure:



All of these steps are repeated every round. When we talk about the round, we means the cluster construction process to the transmission data process from all regions.

3.3.2 Detailed description

Our adaptation of the LEACH protocol proposes improvements to the clustering process. The combination of K-Means clustering algorithm and LEACH improves cluster assignment, increases cluster specificity, and makes the process energy-efficient to prolong the network lifetime.

our proposed kmeans-LEACH protocol which constitutes in two phases.

Kmeans-LEACH is a hierarchical routing protocol, employing a clustering process that divides the network into two levels: cluster-heads and member nodes. The protocol takes place in rounds. Each round consists of two phases: construction and communication. The construction phase is the same for all rounds except the first. Whereas, the communication phase is the same for all rounds.

1. **Construction phase** Our proposal is a combination between the LEACH clustering and Kmeans clustering.

   - In the first round, the construction process follows the kmeans method steps. The nearest nodes are grouped together with a head node that is closest to the center of each partition.

   - In the other rounds, the probabilistic aspect is used for construction. In order to favor nodes that have more energy to become CHs, we calculate the energy ratio of each node. Next, the value of the ratio is assigned to the desired probability of cluster-heads of node associate. In the case of equal probabilities, the node geographically closest to the center of the group is chosen as a cluster leader.

2. **Communication phase**

   The sensors transmit their data to the cluster manager, who aggregates it before transmitting it to the sink. This communication between a cluster-head and the base station is done in a direct way.

Below we introduce the algorithm of the Kmeans-LEACH clustering strategy.

Algorithme 3.1 : Kmeans-LEACH Clustering Algorithm

S': Set of nodes that have not been elected as CH during the previous rounds
R_z: Round number z, where $z \in [1, 2, ..., m]$
$CH(h)$: Node h is a cluster head
$prob[h]$: Desired probability of CH during a round for node h
E_{init}: Initial energy of a node
E_{cons}: Consumed energy by a node
Start

if *first round* R_1 **then**
  Specify the number k of clusters to assign.
  Randomly initialize k centroids
  Repeat Assign each point to its closet centroid.
  Compute the new cetroid (mean) of each cluster.
  Until the centroid position does not change
  Determine the cluster leader geographically closest to the center of cluster
  $CH(h) \leftarrow \{h \in S' | \frac{1}{|P_j|} \sum_{i=1}^{n} S'_i \}$;
else
  for *each round* R_{z+1} **do**
    `// Generate probability values`
    $prob[h] \leftarrow \{h \in S' \mid (E_{init} - E_{cons})/E_{init}\}$;
    `// Choose cluster heads`
    Find the node with maximum $prob[h]$;
    Consider this node as the new cluster head;
    $CH(h) \leftarrow \{h \in S' | \max(prob[h])\}$;
    if *there are multiple nodes with the same probability* **then**
      `// Choose cluster head closest geographically to the group center`
      $CH(h) \leftarrow \{h \in S' | \frac{1}{|P_j|} \sum_{i=1}^{n} S'_i \}$;
    Endif
    `// Build clusters and assign each node to its closet che`
    $arg\, min_P \sum_{j=1}^{k} \sum_{i=1}^{n} \|S_i - C_j\|^2$
  Endfor
Endif

`// Update cluster head:` $MMA[x] = MMA[x] - RDS[h]$
End

3.3.3 Package Format

The information contained in the package is separated into two parts, such as the shows in below:

SindId	CHh	RoundsIdx	S$_{l}ist$	LastRaoundFlag	ProcessingCode	Data

1. **Attributes** The attributes represent the permanent parameters in the package, it consists of:

   SinkId that is used to identify the "sink" base station.

   CHh which is used to identify the group leader concerned by the processing.

   S$_{l}istofnodesthatwerenotelectedCHinpreviousrounds.$

   RoundIdx is the index of the current cycle. The value is initially set at 1' by the Base station in the first cycle.

   LastRoundFlagindicates that the current cycle is the last cycle in all tasks.

2. **Charge** The load includes two types of data. The first is**ProcessingCode** which is used to process detected data. The second is **Data** that carries the data after the aggregation operation.

3.3.4 Sequence Diagram

The general process of our approach is described as:

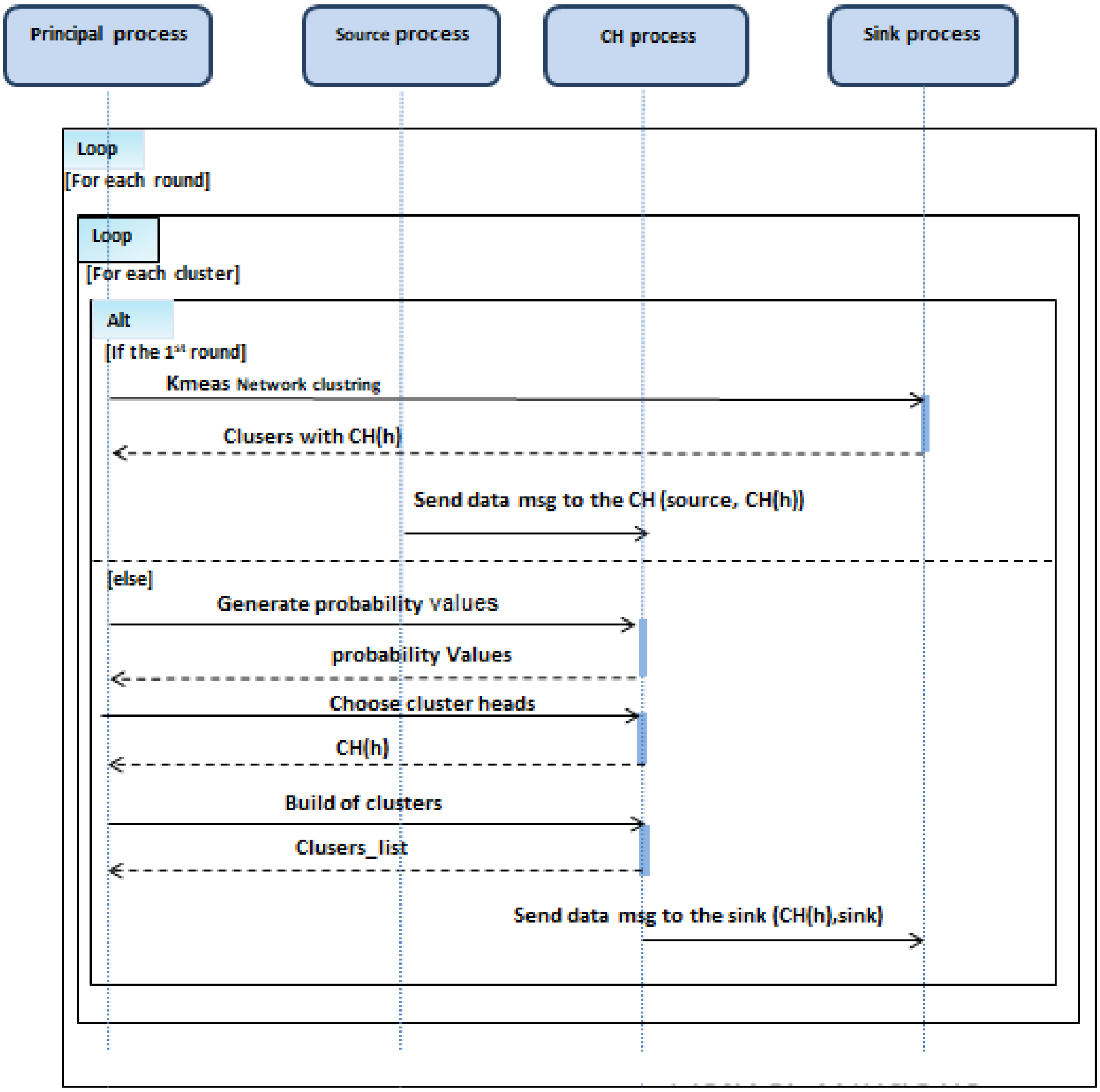
1. We start partitioning the network imported by the base station "Sink" through the Kmeans-LEACH method. where in the first round the nearest nodes are grouped together with a head node for each partition. Whereas, in the other rounds, the probabilistic aspect is used to determine partitioning.

2. Once the leader of each cluster is determined, the sensors transmit their data to the cluster leader, who aggregates it and sends it to the base station for each round.

3. At the end of the current round, the leader of each group calculates the desired cluster leader probability and determines the cluster leader for the next round.

4. Then, the leader of each cluster informs their member nodes, as well as the base station by the new leader.

5. Finally, a new round will be launched by performing the above steps again.



3.4 Conclusion

In this chapter, a combination of the LEACH protocol and kmeans has been proposed to overcome the disadvantages of LEACH and kmeans.

Our LEACH proposal specifically concerns the choice of cluster leaders.

Instead of using the choice of LEACH, we will use the probability calculation of each node to choose as a cluster leader.

After the choice of leader, k-means clustering is a method of that gives a better partition of clusters. In the next chapter, the proposal will be evaluated with a MATLAB simulation.

Chapter 4

Implantation

4.1 Introduction

This chapter unveils the outcomes of the simulations conducted using our implemented K-means-LEACH approach to address the energy issue in WSN. Additionally, it encompasses a comprehensive analysis and validation of the results obtained. Furthermore, we provide an overview of the development tools and platforms employed for implementing the various system components. The chapter culminates in the presentation of the overall results achieved by each algorithm. Finally, discussions are provided, shedding light on the significance and implications of the obtained results.

4.2 Simulation Tools

To implement our K-means-LEACH approach, we used the following tools:

4.2.1 Software Tools

MATLAB® seamlessly integrates a desktop environment optimized for iterative analysis and design procedures with a programming language that enables the direct expression of matrix and array mathematics. This comprehensive platform encompasses the Live Editor, which facilitates the creation of executable notebooks by seamlessly blending code, output, and formatted text. [64]

Figure 4.1: MATLAB Icon

4.2.2 Hardware tools

We used "HP" laptop with the following characteristics:

Processor:Intel(R) Core(TM) i5-6200U CPU @ 2.30GHz 2.40 GHz

RAM:8.00 GB (7.88 GB usable)

System type:64-bit operating system, x64 processor

Windows 10: Microsoft Operating System

4.3 Energy model

The primary factors contributing to the energy consumption of a sensor node are the operations of data capture, processing, and communication. The power consumption, denoted as Ec, can be defined as follows:(hachouf chaima) Ec = Es/sensing+Es/processing+Es/communication where: - Es/sensing: the energy consumption of the sensing unit - Es/processing: the energy consumption of the treatment unit - Es/communication: the energy consumption of the communication unit.

It is equal to the sum of two values:

ETX which is the transmission of energy and ERX which is the reception of energy

Es/communication = ETX + ERX(2) where:

ETX (k, d) = (Eelec * k)+(Eamp *k * d)

ERX (k) = Eelec * k(4)

K: packet size (bits)

d: the distance between the transmitter and the receiver

Eelec: energy to operate the transmitter or receiver circuits

Eamp: transmission amplifier

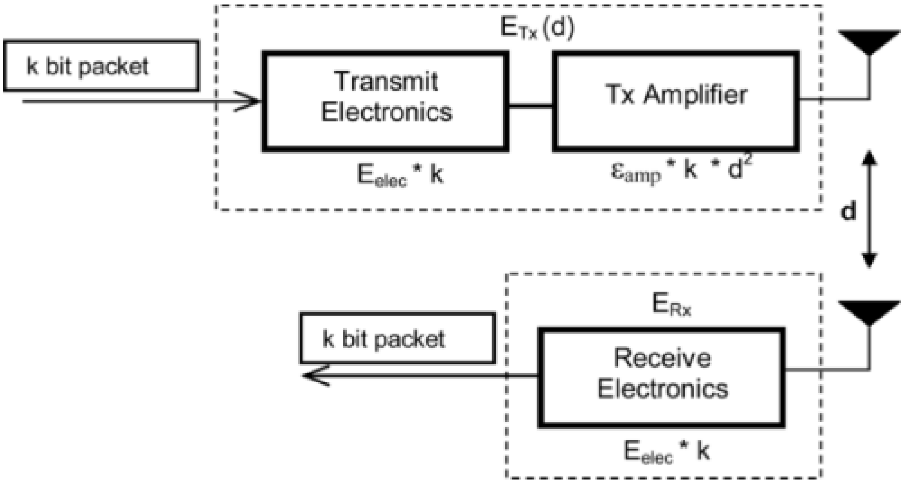


Figure 4.2: Energy Model

4.4 Network model

The network dimension is 100m x 100m then, we deployed the sensor nodes by the matrix, the base station (sink) is located in (50,50). all nodes in the simulated network start with an initial power equal to E0 = 0.5 J and an unlimited amount of data to be sent to the base station. In addition, the power of the base station is unlimited. As shown in the figure below:

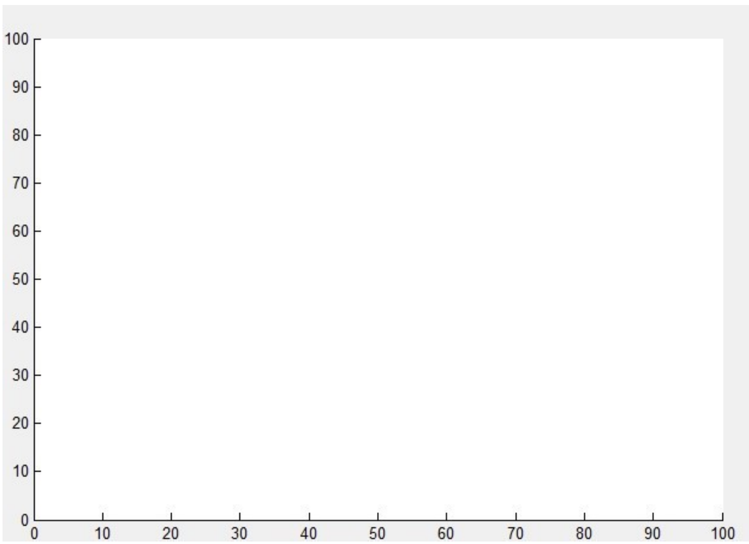


Figure 4.3: Network model

4.5 Program Structure

The user interface contains Node Deployment, K-means and LEACH clustering, and a new "K-means-LEACH" approach.

4.5.1 Program interface

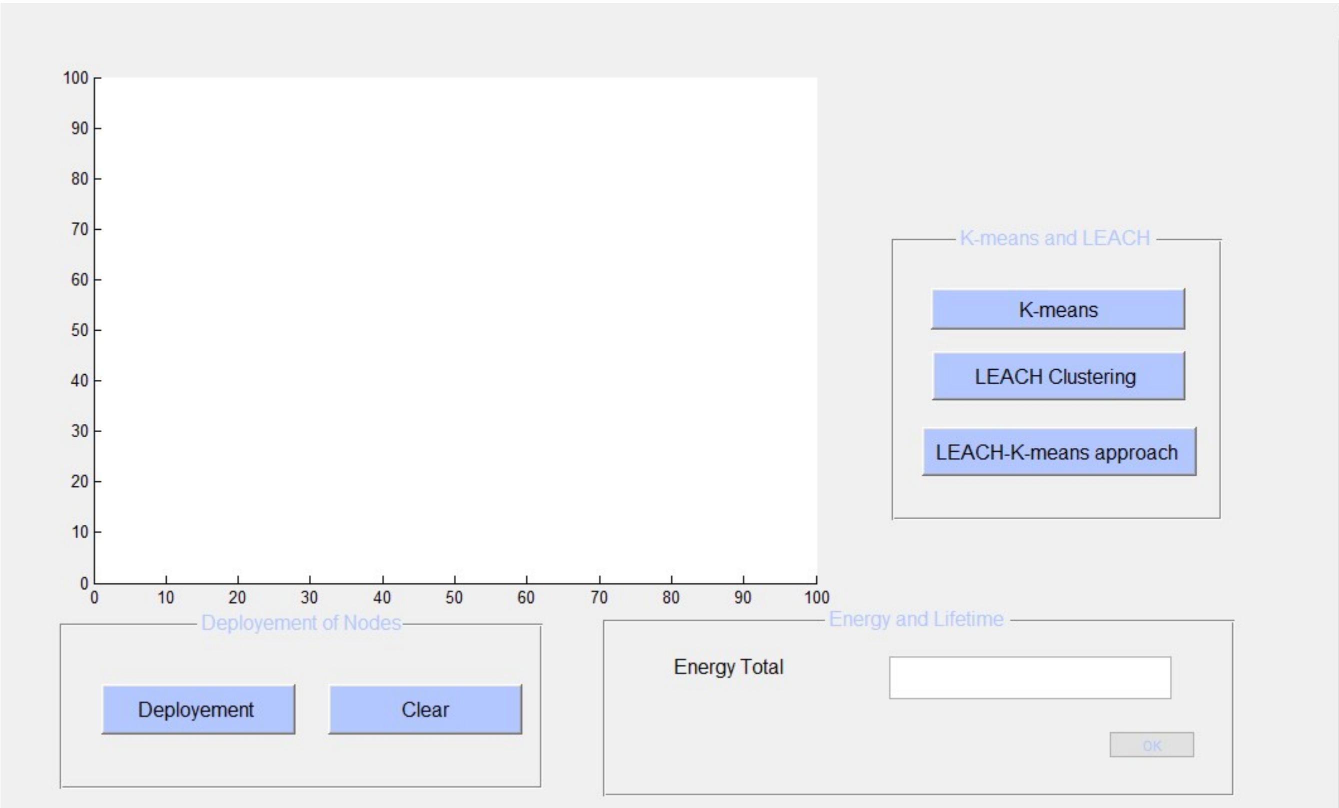


Figure 4.4: Program Interface

4.5.2 Deployment Nodes

This figure show the Deployment of nodes.

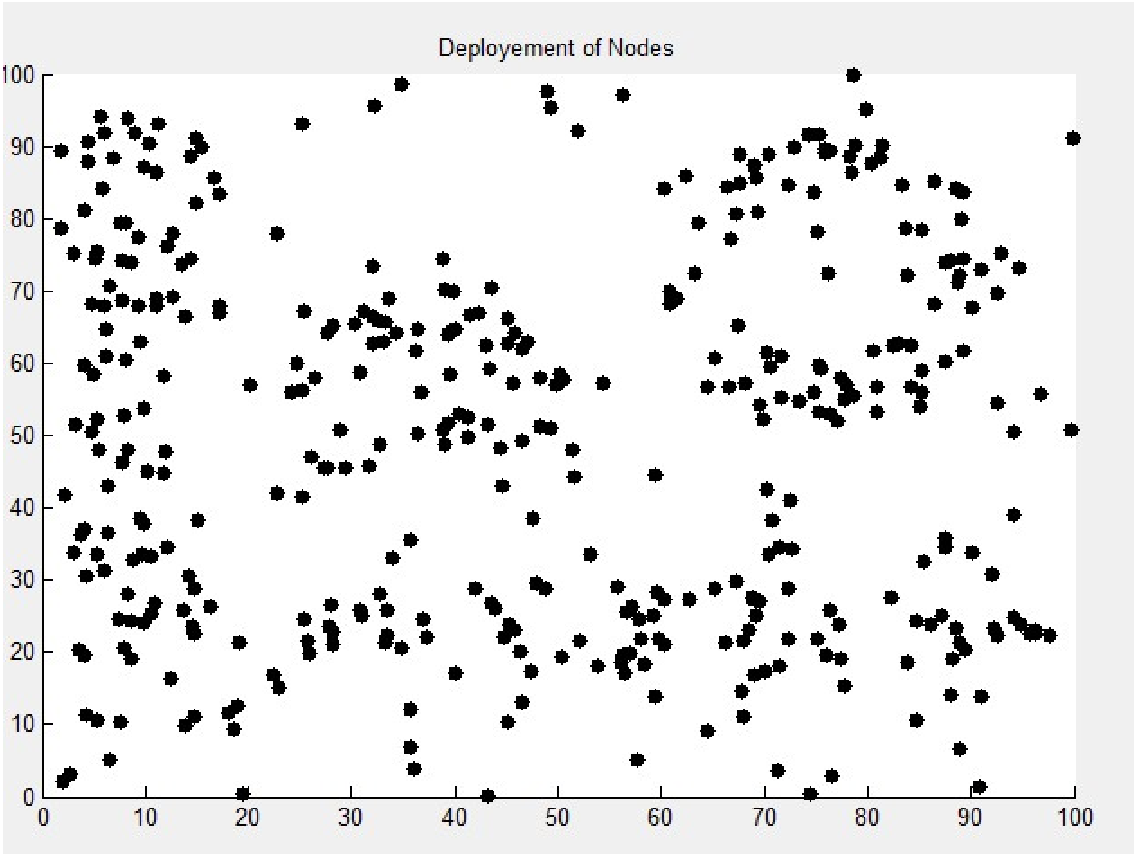


Figure 4.5: Deployment of nodes

4.5.3 K-means Clustering

The figure illustrates the results obtained from applying the k-means algorithm to the data-set. Each data point is assigned to one of the clusters, represented by different colors.

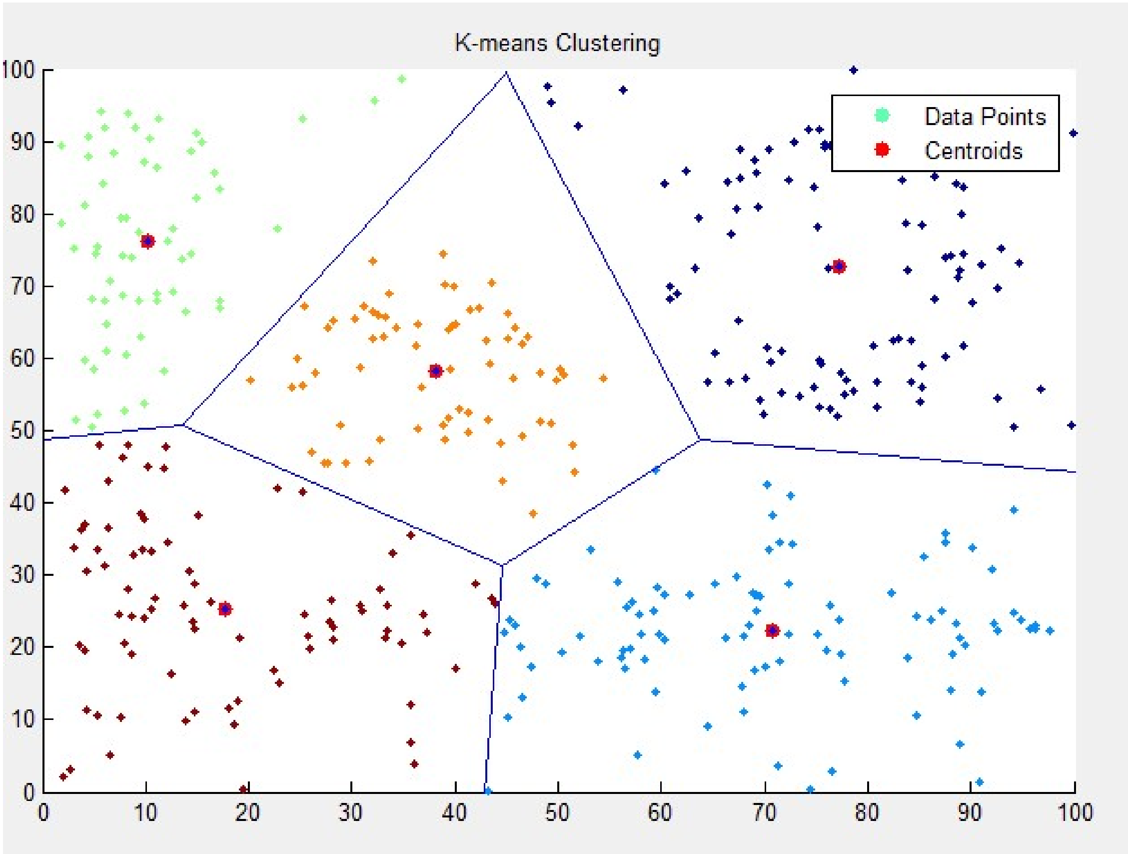


Figure 4.6: K-means Clustering

4.5.4 LEACH Clustering

The figure showcases the results obtained from the LEACH (Low-Energy Adaptive Clustering Hierarchy) algorithm applied to the data-set. The figure provides a visual representation of the clusters formed by the algorithm, showcasing how LEACH effectively organizes the data points into distinct groups based on their energy consumption and proximity to cluster heads. This visualization assists in evaluating the performance and effectiveness of LEACH in optimizing energy usage and prolonging the network lifetime.

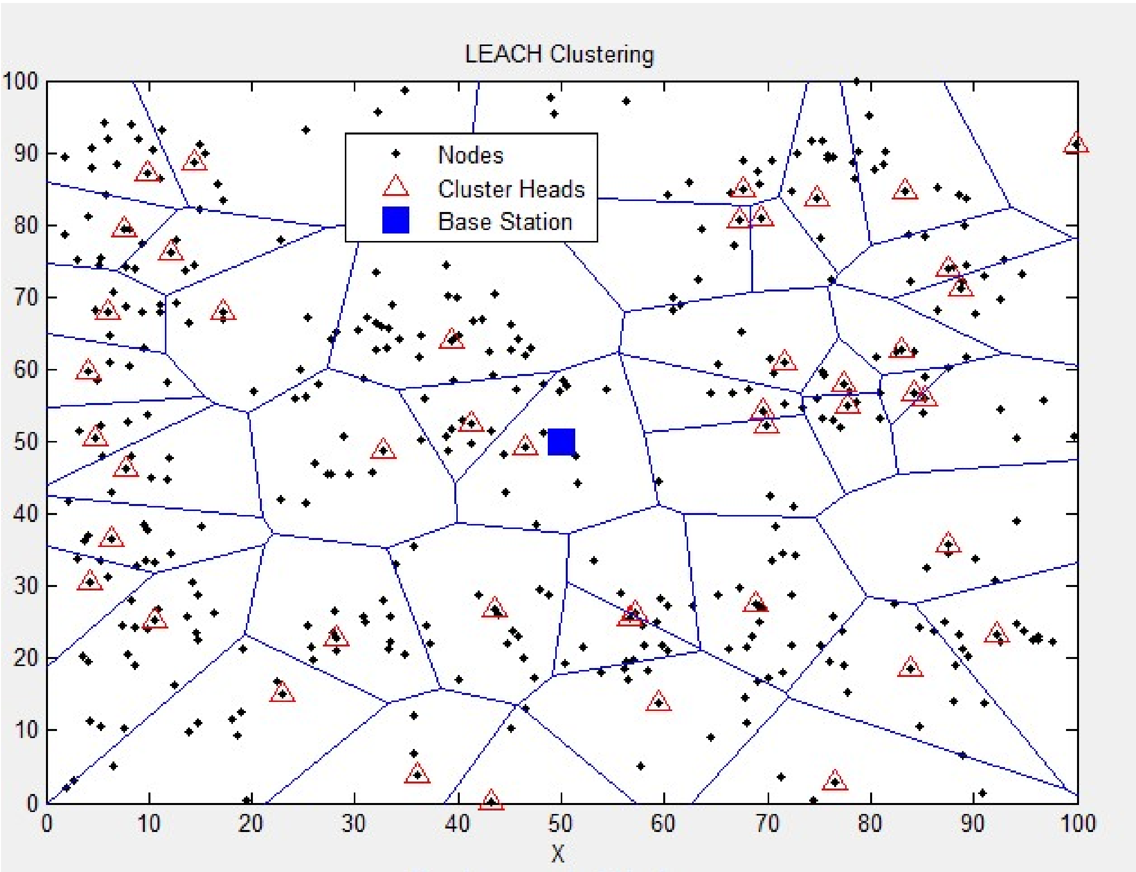


Figure 4.7: LEACH Clustering

4.5.5 K-means-LEACH Approach

The figure illustrates the results obtained from applying the k-means-LEACH Approach to the data-set.

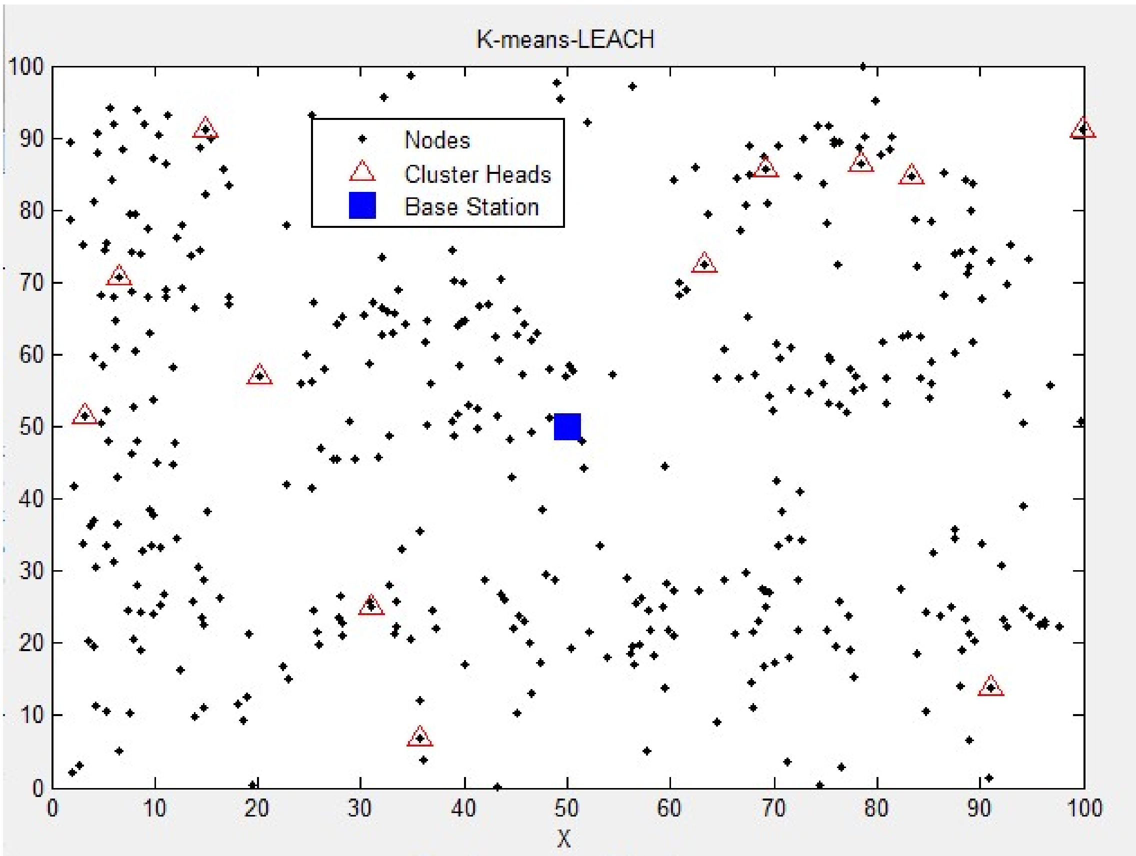


Figure 4.8: K-means-LEACH Clustering

4.6 Simulation

The simulation utilized a set of parameters, which are outlined in the accompanying tables:

4.6.1 Simulation parameters

Network Parameters	
Parameters	value
Size of the network	100m x 100m
Deployment of the nodes	Defines in matrix «matrix.mat»
Size of the packet	1000bits
Range of radio transmission	20m

4.6.2 Node parameters

Node Parameters	
Parameters	value
Number of nodes	400
Maximum number of rounds	30
Base Station ID	401
Node IDs	[1-400]

4.6.3 Energy parameters

Energy Parameters	
Parameters	value
Probability of a node to become CH (LEACH)	0.1
The initial energy of each node	0.5 Joul
The initial energy	200 Joul
aggregation energy	50e-9 Joules/bit
amplification energy	100e-12 Joules/bit/m
transmission energy	50e-9 Joules/bit
Receiving energy	50e-9 Joules/bit

4.7 Results and discussion

In this section, we present the results obtained from running K-means, LEACH, and our k-means-LEACH proposal. Where we discuss and analyze these results gained by comparing the performance measures above.

4.7.1 Energy consumption

Figures (4.9, 4.10) illustrate that the K-means-LEACH consumes less energy than K-means and LEACH protocol, and that means the network lifetime of K-means-LEACH is much bigger than the lifetime of K-means and LEACH because the number of dead nodes is less than the other protocols.

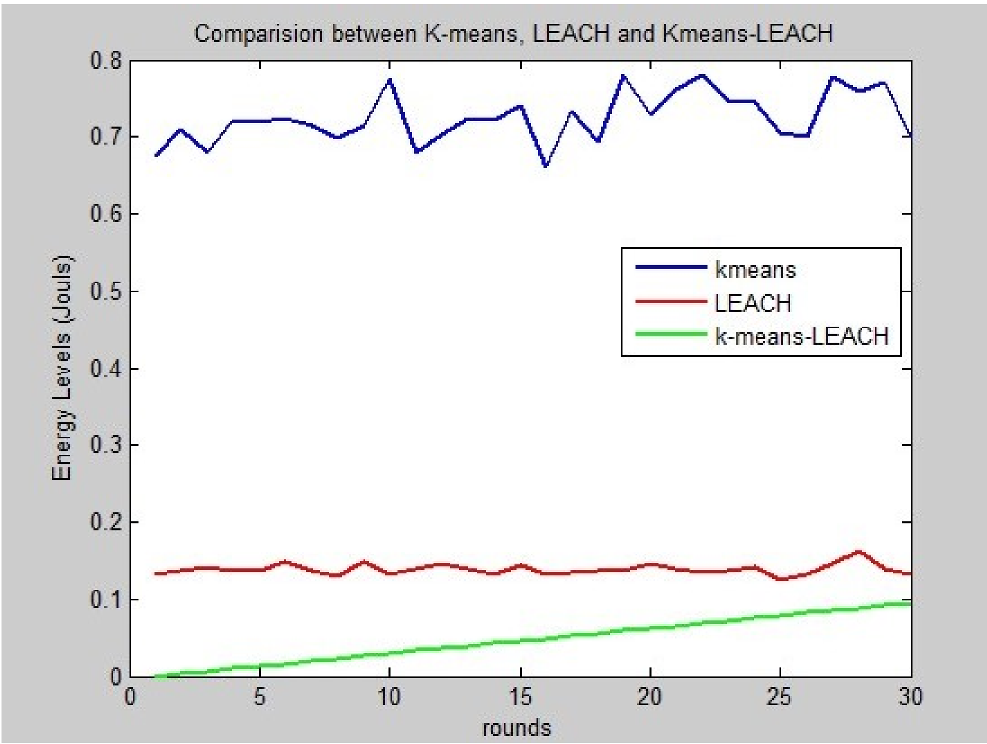


Figure 4.9: Energy consumption for K-means, LEACH, and K-means-LEACH

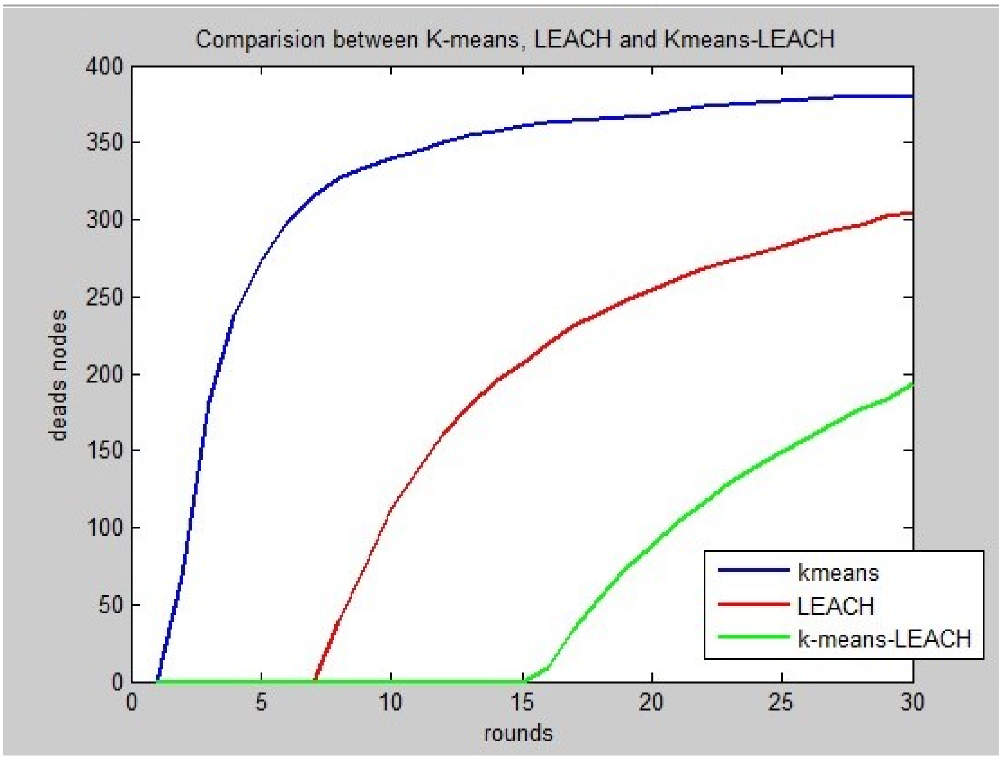


Figure 4.10: Dead Nodes for K-means, LEACH, and K-means-LEACH

4.7.2 processing time

When considering the duration of the task (the processing time), it is evident that K-means-LEACH requires more time compared to both K-means and LEACH. This disparity arises from the inclusion of additional instructions in the K-means-LEACH protocol. figure 4.11 illustrates the difference between them

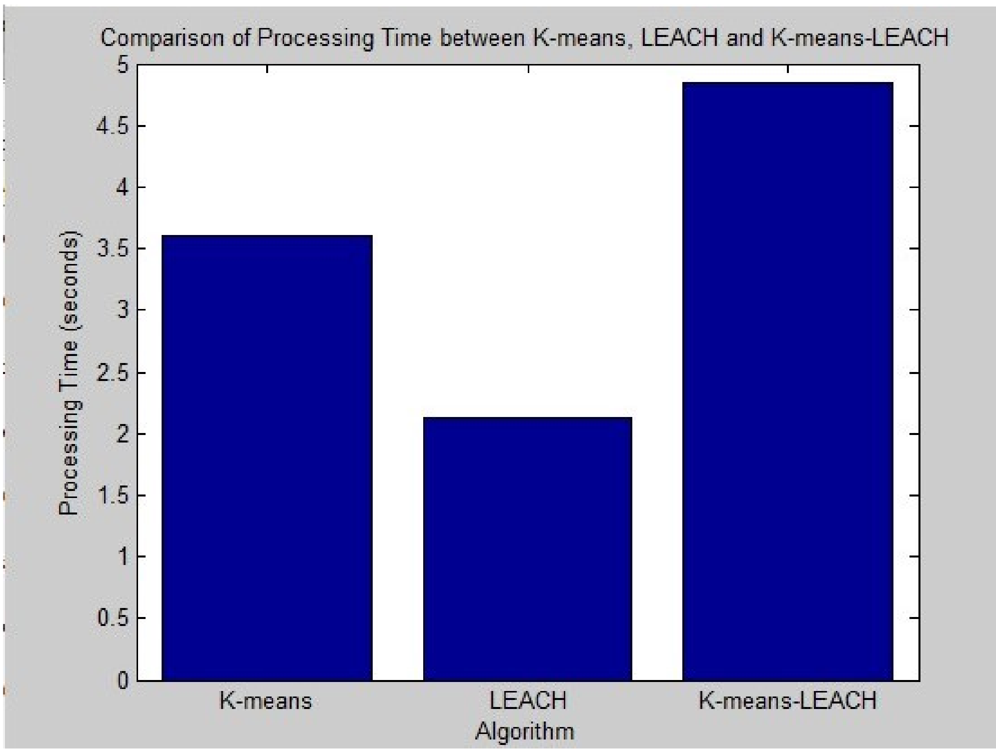


Figure 4.11: Processing Time for K-means, LEACH, and K-means-LEACH

4.8 Conclusion

This chapter explored simulation tools, parameters, and network models. We also discussed comparing three protocols, namely LEACH, K-means, and K-means-LEACH and presented the obtained results. The findings demonstrated that K-means-LEACH outperformed both K-means and LEACH regarding energy consumption and task duration. Consequently, K-means-LEACH succeeded in maximizing the network lifetime, indicating its effectiveness in enhancing overall system performance.

General conclusion

In our research, we focused on exploring the domain of Wireless Sensor Networks (WSN) and its various applications, as well as the challenges that need to be knowing within this field. Specifically, we delved into hierarchical routing protocols and thoroughly examined the functionalities of two well-known protocols: LEACH and K-means.

Moreover, we aimed to enhance the existing LEACH protocol by introducing an adaptation at the phase level. This modification involved the build cluster like k-means and selecting the cluster leader based on the probability of each node, which was determined by its energy level. By incorporating this probability factor, we sought to optimize the protocol's performance. To evaluate k-means-LEACH effectiveness, we conducted simulations and compared the results with those obtained from running both the LEACH protocol and the original K-means protocol.

The simulation outcomes revealed that the K-means-LEACH protocol, with our proposed adaptation, yielded results . Notably, it exhibited significantly lower energy consumption compared to the other protocols. As a consequence, the network's lifetime was extended, which is a crucial factor in WSNs. These findings indicate that our synthesized protocol could potentially offer substantial benefits in terms of network efficiency and longevity.

Bibliography

[1] ADOUANE Sofiane. *Conception d'un protocole de routage hiérarchique efficace en énergie pour les réseaux de capteurs sans fil*. PhD thesis, Université Mohamed Khider – BISKRA, 2021.

[2] Dr.S.Omkumar. Course Material for Wireless Sensor Networks. 2016.

[3] Felicia Engmann, Ferdinand Katsriku, Jamal-Deen Abdulai, and Kofi Adu-Manu. Reducing the energy budget in wsn using time series models. *Wireless Communications and Mobile Computing*, 08 2020.

[4] Lebidi Wafa. *Simulation du comportement du protocole ODD « Overview, Design Concepts, and Details »dans les réseaux de capteurs sans fils avec le simulateur NET-LOGO*. PhD thesis, 2021.

[5] Abdelhak Gueroui, Blaise Yenke, Ado Ari, and Nabila Labraoui. Clustering algorithm for wireless sensor networks : the honeybee swarms nest-sites selection process based approach. *International Journal of Sensor Networks*, 1:1, 01 2016.

[6] Adil Fahad, Najlaa Alshatri, Zahir Tari, Abdullah Alamri, Ibrahim Khalil, Albert Y. Zomaya, Sebti Foufou, and Abdelaziz Bouras. A survey of clustering algorithms for big data: Taxonomy and empirical analysis. *IEEE Transactions on Emerging Topics in Computing*, 2(3):267–279, 2014.

[7] Rezoan Ahmed Nazib and Sangman Moh. Sink-type-dependent data-gathering frameworks in wireless sensor networks: A comparative study. *Sensors*, 21(8):2829, 2021.

[8] Abdullah Al-Odienat, Banan Malahmeh, and Awj Tarawneh. Improved leach protocol for increasing network lifetime based on circular patches clustering. *Universal Journal of Applied Science*, 7, 05 2020.

[9] BOUAMARA Ouarda and BOUAMARA Rosa. *Simulation d'un Réseau de capteur sans fil sous contiki"La localisation d'un mobile"*. PhD thesis, 2016.

[10] Madiou Nabila and Merbouti Kahina. *Approches du clustering dans les réseaux de capteurs sans fil*. PhD thesis, Université Mouloud Mammeri, 2017.

[11] Guemmadi Sabrina. *Adaptation de clustering phase de protocole de routage « LEACH » pour les réseaux de capteurs sans fil*. PhD thesis, 2020.

[12] Basic Introduction of Wireless Sensor Network. Technical report, 2019.

[13] BOUBEKRI Abdelmalek and CHIKOU Youghorta. *Conception et Implémentation d'un Simulateur pour les Réseaux de Capteurs Sans Fil Présenté*. PhD thesis, 2016.

[14] Fabian Nack. An Overview on Wireless Sensor Networks. Technical report.

[15] Mounia GUERTI and Assia HALILEM. *Améliorer les performances des RCSFs à l'aide de protocole de routage basé sur le clustering*. PhD thesis, 2020.

[16] Idres Louiza. *Système de détection d'intrusion hybride et hiérarchique pour les réseaux de capteur sans fil*. PhD thesis, Mouloud Mammeri de Tizi-Ouzou, 2012.

[17] Aina Randrianarisaina. *Modélisation de la consommation d'énergie En vue de la conception conjointe (matériel/logiciel) des applications embarquées. Application aux réseaux de capteurs sans fil (wsn)*. PhD thesis, université de Nantes, 2015.

[18] Shamneesh Sharma, Dinesh Kumar, and Keshav Kishore. Wireless sensor networks-a review on topologies and node architecture. *International Journal of Computer Sciences and Engineering*, 2013.

[19] Muhammad R Ahmed, Xu Huang, Dharmandra Sharma, and Hongyan Cui. Wireless sensor network: characteristics and architectures. *International Journal of Information and Communication Engineering*, 6(12):1398–1401, 2012.

[20] Hiba Bazzi, Ali Haidar, and Ahmad Bilal. Classification of routing protocols in wireless sensor network. pages 1–5, 01 2015.

[21] Optimisation de la durée de vie dans les réseaux de capteurs sans fil. `https://www.rapport-gratuit.com/optimisation-de-la-duree-de-vie-dans-les-reseaux-de-capteurs-sans-fil/`, note = Accessed: 2023-03-20.

[22] Mekki Kada and Nabil Mohammedi. PhD thesis, 2018.

[23] kahla ahmed Tedjani. *planification d'itinéraires basée sur l'algorithme GRASP dans les réseaux de capteurs sans fil*. PhD thesis, 2020.

[24] Gherbi Chirihane. *Algorithme de routage pour les réseaux de capteurs avec prise en charge de la consommation d'énergie.* PhD thesis, 2017.

[25] S Vassilaras. An Overview of Wireless Sensor Networks , their Applications and Technical Challenges. *International Journal of Soft Computing and Engineering (IJSCE), ISSN: 2231-2307*, (2):538–540, 2012.

[26] Naili Bilel and Lamraoui Ahmed. *Mémoire Les Services actifs pour le multicast fiable dans les réseaux de capteurs sans fil.* PhD thesis, 2016.

[27] Mizane Walid. *Contribution à l'amélioration de la consommation d'énergie dans les réseaux de capteurs sans fils (WSN).* PhD thesis, 2021.

[28] Team Tesca. What is wireless sensor network, and types of wsn?

[29] Chaima Hachouf. *Conception d'un système de gestion de clés pour les réseaux de capteur sans fil .* PhD thesis, 2022.

[30] CHEMOUN Karima. *Proposition et implémentation d'un protocole d'agrégation de données basé sur l'approche de clustering dans les RCSFs.* PhD thesis, 2013.

[31] Bounar Seyyid-ali and Kriket Farhat. *Etude du clustering dans les r6seaux de capteurs sans fil.* PhD thesis, 2016.

[32] Sudhir Kumar Gupta, Ashish Ranjan, and Gagandeep Sharma. Clustering technique for Wireless Sensor Network. *Proceedings on 2015 1st International Conference on Next Generation Computing Technologies, NGCT 2015*, pages 223–227, 2016.

[33] Amit Karmaker. *AN ENERGY EFFICIENT AND BALANCED CLUSTERING APPROACH FOR IMPROVING THROUGHPUT OF WIRELESS SENSOR NETWORKS.* PhD thesis, 2017.

[34] Vivek Katiyar, Narottam Chand, and Surender Soni. A Survey on Clustering Algorithms for Heterogeneous Wireless Sensor Networks. 2010.

[35] Menhane Chihab Eddine and Aissa Abdelouahab. *Réseaux de capteurs sans fil : Mesure de l'efficacité énergétique des techniques de clustering.* PhD thesis, 2019.

[36] BOULARAS Soufiane and DEHBINE Youcef. *ROUTAGE HIERARCHIQUE SUR LES RESEAUX DE CAPTEURS SANS FIL.* PhD thesis, 2014.

[37] SA Elavarasi, J Akilandeswari, and B Sathiyabhama. A survey on partition clustering algorithms. *Learning*, 1(1):1–14, 2011.

[38] Bessaddoug Abdallah and Boudjemai Ilyes. *clustering et routage dans les reseaux de capteurs corporels sans fils*. PhD thesis, 2021.

[39] Sakshi Patel, Shivani Sihmar, and Aman Jatain. A study of hierarchical clustering algorithms. *2015 International Conference on Computing for Sustainable Global Development, INDIACom 2015*, 3(10):537–541, 2015.

[40] Kamalpreet Bindra and Anuranjan Mishra. A Detailed Study of Clustering Algorithms. *International Conference on Current Trends in Computer, Electrical, Electronics and Communication, CTCEEC 2017*, pages 752–757, 2018.

[41] Pradeep Rai and Shubha Singh. A Survey of Clustering Techniques. *International Journal of Computer Applications*, 7(12):1–5, 2010.

[42] DEBBAL Bouchra and BENDIMERAD Ahlem. *Une nouvelle approche de clustering pour améliorer les performances des réseaux de capteurs sans fil*. PhD thesis, 2019.

[43] Neeraj Jain, P Sinha, and SK Gupta. Clustering protocols in wireless sensor networks: A survey. *International Journal of Applied Information System (IJAIS)*, 5(2), 2013.

[44] Xuxun Liu and Jinglun Shi. Clustering routing algorithms in wireless sensor networks: an overview. *KSII Transactions on Internet and Information Systems (TIIS)*, 6(7):1735–1755, 2012.

[45] Riham Elhabyan. *Clustering and routing protocols for wireless sensor networks: design and performance evaluation*. PhD thesis, Université d'Ottawa/University of Ottawa, 2015.

[46] Rajendran Sugin Elankavi and R. Udayakumar. A review on clustering algorithms in wireless sensor networks. 2013.

[47] Mohammed Al-Shalabi, Mohammed Anbar, Tat-Chee Wan, and Ahmad Khasawneh. Variants of the low-energy adaptive clustering hierarchy protocol: Survey, issues and challenges. *Electronics*, 7(8):136, 2018.

[48] Chihab Eddine MENHANE, Aissa ABDELOUAHAB, and Messaoud Encadreur KARA. *Réseaux de capteurs sans fil: Mesure de l'efficacité énergétique des techniques de clustering*. PhD thesis, Université jijel, 2019.

[49] Shahrzad Dehghani, Mohammad Pourzaferani, and Behrang Barekatain. Comparison on energy-efficient cluster based routing algorithms in wireless sensor network. *Procedia Computer Science*, 72:535–542, 12 2015.

[50] Zaib Ullah, Leonardo Mostarda, Roberto Gagliardi, Diletta Cacciagrano, and Flavio Corradini. A comparison of heed based clustering algorithms–introducing er-heed. In *2016 IEEE 30th international conference on advanced information networking and applications (AINA)*, pages 339–345. IEEE, 2016.

[51] Xuxun Liu and Jinglun Shi. Clustering Routing Algorithms In Wireless Sensor Networks : An Overview Clustering Routing Algorithms In Wireless Sensor Networks : An Overview. 2016.

[52] Abdul Rehman Khan, Nitin Rakesh, Abhay Bansal, and Dev Kumar Chaudhary. Comparative study of wsn protocols (leach, pegasis and teen). In *2015 Third International Conference on Image Information Processing (ICIIP)*, pages 422–427. IEEE, 2015.

[53] Haydar Marhoon, Massudi Mahmuddin, and Shahrudin Awang Nor. Chain-based routing protocols in wireless sensor networks: A survey. *ARPN Journal of Engineering and Applied Sciences*, 10:1389–1398, 02 2015.

[54] Nasrin Azizi, Jaber Karimpour, and Farid Seifi. Hcte: Hierarchical clustering based routing algorithm with applying the two cluster heads in each cluster for energy balancing in wsn. *International Journal of Computer Science Issues*, 9, 01 2012.

[55] Chavan Pragnesh H. CLUSTERING IN WSNs : A REVIEW. 3(9):2149–2153, 2016.

[56] Hafsa Ennajari, Yann Ben Maissa, and Salma Mouline. Energy efficient in-network aggregation algorithms in wireless sensor networks: A survey. *Lecture Notes in Electrical Engineering*, 397:135–148, 2017.

[57] Lu Li-fang and Cheng chew Lim. Weight-based clustering routing protocol for wireless sensor networks. pages 554 – 559, 09 2009.

[58] Mervat Mustafa Raouf. Clustering in Wireless Sensor Networks (WSNs). (57), 2019.

[59] Moufida Maimour, Houda Zeghilet, and Francis Lepage. *Cluster-based Routing Protocols for Energy-Efficiency in Wireless Sensor Networks*. 12 2010.

[60] Sariga Arjunan and Sujatha Pothula. A survey on unequal clustering protocols in wireless sensor networks. *Journal of King Saud University - Computer and Information Sciences*, 31(3):304–317, 2019.

[61] Muhammad Rahman. Energy-efficient wireless sensor network with an unequal clustering protocol based on a balanced energy method (eeucb). *Sensors*, 21, 01 2021.

[62] Jawahar Senthil Kumar V and Ram Prabha S. A New Approach of Unequal Clustering in Wireless Sensor Networks. 3(16):1–5, 2015.

[63] A. Manjeshwar and D.P. Agrawal. Teen: a routing protocol for enhanced efficiency in wireless sensor networks. In *Proceedings 15th International Parallel and Distributed Processing Symposium. IPDPS 2001*, April 2001.

[64] MATLAB. https://www.mathworks.com/products/matlab.html?s_tid=hp_ff_p_matlab, 2023.