

**Republic of Iraq
Ministry of Higher Education and
Scientific Research
University of Babylon
College of Science for Women
Department of Computer Science**



Energy-efficient Scheduling Approach Using Clustering and Encoding in Wireless Sensor Networks

A Thesis

Submitted to the Council of College of Science for Women, the
University of Babylon in a Partial Fulfilment of the Requirements for the
Degree of Master in Science\ Computer Sciences

By

Lina Waleed AL-Shammary

Supervised By

Prof. Dr. Ali Kadhum Idrees

October 2021 A. D.

1442 A. H.

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

((قَالُوا سُبْحَانَكَ لَا عِلْمَ لَنَا إِلَّا

مَا عَلَّمْتَنَا

إِنَّكَ أَنْتَ الْعَلِيمُ الْحَكِيمُ))

صَدَقَ اللَّهُ الْعَلِيُّ الْعَظِيمُ

Dedication

I dedicate this thesis to **Allah** Almighty my great creator, my lord, and my strong pillar...

This thesis is dedicated to the spirit of my **Father** who I wished to attend my discussion...

Then I dedicated it to my dear **Mother**, who has always been a source of inspiration, encouragement, and stamina to undertake my...

And finally, to all my brothers and sisters, especially my big brother, my spiritual father, and the source of my strength, **Khaldun**, who was the best brother, father, and support throughout my life.

Supervisor Certification

I certify that project entitled "**Energy-efficient Scheduling Approach Using Clustering and Encoding in Wireless Sensor Networks**" was prepared at the Department of Computer Sciences/ College of Science for the Women/ University of Babylon, by (**Lina Waleed AL-Summary**) as partial fulfilment of the requirements for the degree of Master in Computer Science.

Signature:

Name: Prof. **Dr. Ali Kadhum Idrees**

Date: / / 2022

Address: University of Babylon/College of Science for Women

The Head of the Department Certification

Given the available recommendations, I forward the research entitled “**Energy-efficient Scheduling Approach Using Clustering and Encoding in Wireless Sensor Networks**” for debate by the examination committee.

Signature:

Name: **Dr. Farah Muhammed Hassan.**

Date: / / 2022

Address: University of Babylon/College of Science for Women

Acknowledgments

I would like to express my special thanks of gratitude to my supervisor **Dr. Ali Kadhum Idrees** who gave me this golden opportunity to do this thesis on the topic,

which also helped me in doing a lot of research and I came to know so many things. Without his advice and active support, this project would not have been possible. I am thankful to him.

I offer to **Dr. Ali Kadhum Idrees** all my respect and sincere prayers for his continuous support, motivation, and kind words.

I would also like to acknowledge the University of Babylon – College of Science for Women – Department of Computer Science for providing all sorts of support in conducting this thesis.

Last but not least, I would like to thank all of my family and my friends who helped me a lot in finalizing this thesis within the limited time frame.

Abstract

The Wireless Sensor Networks (WSNs) represent the essential element of many applications in the Internet of Things (IoT) and the smart cities in the current days and Future. The deployment nature of these devices in many applications leads to an increase in the volume of redundant data transferred over the IoT network. This can lead to degradation of the network performance and deplete the limited battery power of the smart devices. Therefore, it is important to deal with this problem by turning off the redundant sensors that transmit the similar data to the sink and activating a smaller number of sensor nodes in the IoT network. This reduces the redundant sensed readings and decreases the overhead of communications, thereby, extending the WSN lifetime.

In this thesis, an Energy-efficient scheduling approach using clustering and encoding in WSNs is proposed. The proposed approach is implemented at two levels: sensor nodes and the gateway. The sensor nodes capture the data and compress them using either Modified Run Length Encoding (MRLE) or Huffman Encoding (HE). At the gateway level, the sensor nodes are scheduled based on the clustering methods to produce the best schedule of sensor nodes that stay active in the next period. The sensor nodes are clustered based on the received vectors of data into groups and one sensor will be elected in each group to stay active in the next period.

The simulation results (Custom simulator-based Python language) are conducted to prove the effectiveness of the proposed technique, where it provides competitive results in comparison with some other work in terms of energy consumption, active sensor ratio, transmitted data ratio, and percentage of lost data. The agglomerative and Modified Run Length

Encoding (AgglomMRLE) decreased respectively the transmitted data and consumed energy from 94.3% up to 98.8% and from 99.5% up to 99.8% compared to the Prefix Frequency Filtering (PFF). The Kmeans and Modified Run Length Encoding (KmeansMRLE) decreased respectively the transmitted data and consumed energy from 75.9% up to 78.7% and from 95.3% up to 98.1% compared to the PFF. The AgglomMRLE decreased the percentage of lost data from 5.3% up to 21.1 % compared to the PFF.

Table of Contents

Abstract	VI
Table of Contents	VIII
List of Tables	XI
List of Algorithms.....	XII
List of Figures	XIII
List of Abbreviations	XIV
List of Publications	XVI
Chapter One: General Introduction.....	1
1.1 Introduction	1
1.2 Problem Statement	2
1.3 Main Objectives of this Thesis.....	3
1.4 Main Contribution of this Thesis	3
1.5 Literature Review	4
1.6 Thesis Organization	12
Chapter Two: Theoretical Background.....	13
2.1 Introduction.....	14
2.2 Wireless Sensor Networks (WSNs).....	14
2.3 WSN Applications	17
<u> 2.3.1</u> Medical Applications.....	18
<u> 2.3.2</u> Military Applications.....	18
<u> 2.3.3</u> Environmental Monitoring	18
<u> 2.3.4</u> Agricultural sector	18
<u> 2.3.5</u> Structural monitoring.....	19

2.4	Challenges in WSNs	19
	2.4.1 Energy	19
	2.4.2 Security	20
	2.4.3 Quality of Service	20
	2.4.4 The Dynamic Topologies	20
	2.4.5 Hardware Limitations	20
2.5	Energy-Efficient Mechanisms in WSNs.....	21
	2.5.1 Schedule Mechanism.....	21
	2.5.2 Clustering.....	24
	A. General CH Selection Methods	27
	B. Objectives of Clustering.....	28
	C. Clustering Algorithms.....	29
	2.5.3 Energy-efficient Routing	30
	2.5.4 Energy Harvesting Techniques.....	31
	2.5.5 Data compression.....	32
	A. Encoding RLE.....	34
	B. Huffman coding	35
2.6	Data Clustering	36
	2.6.1 Partitional Clustering Algorithms.....	36
	2.6.2 Hierarchical Clustering Algorithms.....	37
2.7	Data Correlation.....	38
2.8	Network Lifetime.....	39
2.9	Energy Consumption Model.....	40
2.10	Performance metrics	41

2.11	Summary	43
Chapter Three: The Proposed Approach.....		44
3.1	Introduction	45
3.2	The Framework of the Proposed Work	45
<u>3.2.1</u>	Sensor level.....	47
<u>3.2.2</u>	Gateway level	48
<u>A.</u>	Scheduling based agglomerative Algorithm.....	50
<u>B.</u>	Scheduling based k-means Algorithm	53
3.3	Summary	56
Chapter Four: Simulation Results and Discussions.....		58
4.1	Introduction.....	58
4.2	Simulation Framework	58
4.3	Performance Evaluation and Analysis.....	60
<u>4.3.1</u>	Selecting the Number of Clusters (NC).....	60
<u>4.3.2</u>	Active Sensor Ratio (ASR).....	61
<u>4.3.3</u>	Transmitted Data Ratio (TDR)	63
<u>4.5.4</u>	Energy Consumption (EC)	66
<u>4.5.5</u>	Percentage of Lost Data (PLD)	68
4.5	Summary	70
Chapter Five: Conclusions and Suggestions for Future Works.....		72
5.1	Conclusions.....	72
5.2	Future Works	73
References		75

List of Tables

<i>Table No.</i>	<i>Table Title</i>	<i>Page No.</i>
1.1	Summary of published researches in scheduling in WSNs	8
4.1	Parameters setting	60
4.2	Number of Clusters (NC) vs ASR, TD, EC, and TLM	61

List of Algorithms

<i>Algorithm No.</i>	<i>Algorithm Title</i>	<i>Page No.</i>
2.1	Huffman Encoding	35
2.2	Huffman Decoding	36
3.1	MRLE	48
3.2	MRLD	49
3.3	Scheduling Algorithm based agglomerative clustering	52
3.4	Scheduling Algorithm based k-means Algorithm	53
3.5	Scheduling Algorithm	55

List of Figures

<i>Figure No.</i>	<i>Figure Title</i>	<i>Page No.</i>
2.1	Wireless Sensor Network Architecture	15
2.2	The components of Sensor Node	16
2.3	Wireless Sensor Network Applications	17
2.4	Wake up Scheduling Schemes Classification	23
2.5	Clustering in Wireless Sensor Networks	26
2.6	Energy Harvesting Techniques	32
3.1	Flowchart of Proposed Approach	46
4.1	The deployment of the sensors in the Intel Laboratory	59
4.2	Active Sensor Ratio	63
4.3	Transmitted Data Ratio	65
4.4	Energy Consumption	67
4.5	Percentage of Lost Data	69

List of Abbreviations

ASR	Active Sensor Ratio
ADC	Analog to Digital Converters
AHC	Agglomerative Hierarchical Clustering
AP	Affinity Propagation
ACO	Ant Colony Optimization
ARE	Average of Residual Energy
BS	Base Station
CH	Cluster Head
CADS	Content-based Adaptive and Dynamic Scheduling
CPU	Central Processing Unit
DHC	Divisive Hierarchical Clustering
DRL	Deep Reinforcement Learning
EC	Energy Consumption
E²S-DRL	Energy-Efficient Scheduling using the Deep Reinforcement Learning.
EH	Energy harvesting
EBMS	Energy-Balanced Multisensory Scheduling Strategy
EFMSS	Energy-First Multi-point relays Set Selection
ESCSP	Evolution Strategy-based Cover Scheduling Problem
EASS	Energy-Efficient Adaptive Tabulation Scheme
HE	Huffman Encoding
FDN	First Dead Node
IoT	Internet of Things
JPASR	Joint Power Allocation and Secure Routing
MRLE	Modified Run Length Encoding
MuDiLCO	Multi-round Distributed Lifetime Coverage Optimization
NS	Number of Sensor nodes

NC	Number of Clusters
PFF	Prefix Frequency Filtering
PRBDS	Packet Rank Based Data Scheduling
QoS	Quality-of-Service
RAM	Random-Access Memory
RL-Sleep	Reinforcement Learning- sleep
RSCP	Routing Secure Connection Probability
RW	Random Walks
RPR	Received Packet Ratio
SCCS	Spatial Correlation-based partial Coverage Scheduling
SRE	Standard Deviation of Residual Energy
TD	Transmitted Data
TLM	Total Lost Measures
WSN	Wireless Sensor Network
WSN-CSP	Cover Scheduling Problem in Wireless Sensor Networks
ZbC	Zone-based Clustering

List of Publications

- ❑ Lina Waleed Jawad, and Ali Kadhum Idrees (2021).
Energy-efficient Scheduling based Clustering Approach for Wireless Sensor Networks of IoT, 9th International Conference of applied science and technology (ICAST2021): Karbala, Iraq, publishing in AIP (WOS/Scopus)Proceedings,
<https://aip.scitation.org/apc/info/forthcoming>.
- ❑ Lina Waleed Jawad, and Ali Kadhum Idrees (2021).
Integrated Encoding and Scheduling Protocol for Improving the Lifetime in Wireless Sensor Networks. International Journal of Computer Applications in Technology, (**In press**), (**Clarivate**), Indexed in Web of Science, Inderscience Publishers,
<https://www.inderscience.com/info/ingeneral/forthcoming.php?jcode=IJCAT>
- ❑ Lina Waleed Jawad, and Ali Kadhum Idrees (2021).
Hybrid Energy-efficient Data Processing Protocol in Edge-based IOT Networks, International Journal of Sensor Networks. Submitted (under review).

Chapter One
General Introduction

Chapter One

General Introduction

1.1 Introduction

Low-power WSNs have been introduced as a result of recent advancements in wireless communication technology and the production of inexpensive wireless devices. Due to their ease of deployment and the multi-functionality of the sensor nodes of WSN, it becomes more popular and highly grown in the last years, Therefore, its uses and applications increased [1].

WSNs are networks consisting of a lot of small, isolated sensors that are dispersed across defined regions and communicate with one another via wireless connections [2]. The major role of these sensor nodes is to monitor numerous physical variables in their environment and communicate their values to a network component called a Base Station (BS), that collects all the information to be further processed. This last device is called the "Base Station (BS)", "Gateway", or "Sink" [3].

In general, more sensor nodes are deployed than required (compared with the optimal placement) to implement the suggested task [4]. This dense deployment of nodes in the target area can result in optimal target monitoring and ensure maximum coverage of the area.

On the other hand, data transmission operations to the BS may increase, which leads to drain the sensor nodes' energy. The depletion of node energy does not affect the node only but also affects the overall network lifetime due to sending similar data from more than one sensor

node to the BS. This leads to the early death of the node and this increases the need to exploit the energy-saving strategies.

Several methods have been introduced to save energy and improve sensor batteries' lives while monitoring a specific sensor field such as routing [5], clustering [6], aggregation [7], schedules [8], and data compression [9], etc.

One promising solution suggested by this thesis is to provide periodic sleep scheduling for sensors, which means keeping part of the sensors in the sleeping state and the other in the active state. In addition, the main objective of sensor activity scheduling is to decrease the number of active sensor nodes; therefore, the power consumption will be saved in a balanced way and the transmission of similar data from different nodes will be significantly reduced while maintaining an acceptable level of required performance.

1.2 Problem Statement

In WSN, the high-density deployment of sensor nodes in the area to be monitored led to an increase in the redundant data transmitted to the sink node and caused unnecessary power consumption since many nodes remain active and sense the same data. In addition, the sensor nodes are usually powered only by batteries, and it is infeasible and costly to replace once sensor nodes have been deployed, so energy efficiency is always a primary concern in WSNs [10]. Therefore, we need a method for scheduling the sensor activities (sleep/active) in an energy-efficient way in WSN to minimize the redundant in the data transmitted to the sink while maintaining an acceptable level of data accuracy, and eventually, save the

power of the sensor nodes since the power is the most important constraint in the sensor nodes that must consider when designing WSN.

1.3 Main Objectives of this Thesis

1. Propose an energy-efficient clustering approach in WSNs.
2. Integrate the scheduling approach with data encoding to further reduce the transmitted data towards the sink.
3. Extensive simulations based on the selected network simulator and by using real measures.
4. Comparison with other works and publications is one of our goals.

1.4 Main Contribution of this Thesis

The primary contributions of this thesis are:

1. Suggests an Energy-efficient scheduling approach using clustering and encoding in WSNs for decreasing the transmission cost and saving energy thus extending the lifetime in the WSNs.
2. Propose Scheduling Algorithm based agglomerative Algorithm that integrated with **MRLE** or **HE** encoding to minimize the size of sent data to the gateway and preserve the power of sensor batteries thus extending the IoT network lifetime.
3. Propose Scheduling Algorithm-based k-means Algorithm that integrated with **MRLE** or **HE** encoding to exploit the temporal and spatial correlation between the sensor nodes and remove the redundant data using compression and scheduling the sensor nodes. The sensor nodes are clustered based on the received vectors of data into groups and one sensor will be elected in each group to stay active in the next period.

4. Several experiments using custom simulator-based python programming language are conducted using real data from the sensors deployed in Intel Lab. To show the superiority of the proposed technique in comparison with one method named **PPF** [11]. The results based on various measures like the active sensor nodes ratio, transmitted data ratio, energy consumption, and percentage of lost data illuminate that the suggested technique can decrease the energy usage, and transmitted data with a suitable number of active sensor nodes per period.

1.5 Literature Review

The main aim of this review is to thoroughly examine published works of literature on employing scheduling techniques to extend the lifetime of WSNs like in {[11] [12], [13], [14], [15], [16], [17], [18], [19] , and [20]} .

Bahi, J. M. et al. in (2014), proposed the approach for data aggregation in sensor networks. Our data aggregation method works in two phases, the first one at the nodes level, which we call local aggregation and the second at the aggregators level. At each period p each node sends its aggregated data set to its proper aggregator which subsequently aggregates all data sets coming from different sensor nodes and sends them to the sink. used the Jaccard similarity function to estimate the similarity between sets of data measures [11].

Feng, J. and Zhao, H. (2018), proposed a distributed and Energy-Balanced Multisensory Scheduling Strategy (EBMS). This considers the scheduling of sensing and communication modules for each one of the nodes in the target tracking WSNs. The network in "EBMS" is orderly into

clustering structures. Adaptively, a sleep time is assigned by each one of the cluster heads to its cluster members based on their position. EBMS can also be used to balance the energy consumption of sensor nodes by adopting an energy-balanced measurement metric in the procedure of multisensory node scheduling [12].

Mhatre, K. P. and Khot, U. P. (2020), developed an ENS_OR algorithm with sleep mode to reduce energy wastage due to passive listening regarding the nodes in ENS_OR. The suggested approach puts nodes that are not involved in packet data relaying to sleep or turns them on. The suggested sleep algorithm is intended to function in conjunction with opportunistic routing to enhance the ENS OR's energy conservation and network performance [13].

Banerjee, P. S. et al. in (2020), proposed a temperature adaptive intelligent sleep-scheduling strategy (RL-Sleep) for the wireless sensor nodes. This approach depends on Reinforcement Learning that allows a node in a network to detect its surroundings and decide autonomously on an activity (sleep, listen, transmit) that is beneficial to the network's stability. As a result, the RL-SLEEP shields the sensor nodes from power exhaustion and allows them to make a state transition in response to changes in the ambient temperature. It also creates intermediate states between sleeping and active modes, allowing sensor nodes to move between them intelligently [14].

Khan, M. N. et al. (2020), A Content-based Adaptive and Dynamic Scheduling (CADS) employing two ways communication model in WSNs is proposed as a novel approach. CADS adopts a new state depending on contents regarding detected data packets and dynamically adjusts a node state during data aggregation to eliminate duplicated data values and limit

forwarding of superfluous data packets. CADS is used to avoid redundant data values and reduce the forwarding of unnecessary data packets. Four states (Active-Live, Sleep-Live, Active-Sleep, Sleep-Sleep) have been defined and controlled by using control messages transmitted by the analyzer module of BS [15].

Feng, W. et al. (2020), a joint power allocation and secure routing strategy (JPASR) is suggested for optimizing the routing secure connection probability (RSCP) under power restrictions. The secure uplink transmission and energy-saving might be ensured when combined with a level-by-level sleeping scheduling approach. In the downlink, the energy-first multipoint relays set selection mechanism (EFMSS) was utilized to choose the backbone nodes that will broadcast messages, while backbone nodes are woken up using an identical level-by-level sleep scheduling technique as in uplink. The key occurrences are handled suitably in the two-step procedure, while the responses are disseminated to the whole network [16].

Khan, M. N. et al., in (2020), for addressing the energy efficiency problem of cloud energy in embedded networks. The authors proposed an Energy-Efficient Adaptive Tabulation Scheme (EASS), in Wireless Sensor Mesh Networks. A sensor node configures and schedules its functions/roles according to the contents of sensed data packets and frequency of generated traffic. Four states (Probe/Listen/Sleep/Active) were specified depending on internal components regarding the sensor node. The contents related to the sensed data packets have been used for configuring the four-state model. When the contents of detected data packets stay the same for three consecutive inputs, the node state is changed. The contents/pattern of sensed data packets determine which of

such 4 states is selected. Energy consumption is reduced and energy efficiency is improved through state-based sensing [17].

Yang, G. et al. in (2020), in wireless body area networks, an energy-efficient approach for scheduling and routing is presented. The number of broadcast flooding is lowered depending on the distance between CN, and nodes CN during the protocol's initialization stage, reducing the time required for collecting network information. Energy-balanced routings are chosen at the routing setup phase by taking into account the channel condition, node location, and node type. The time slot is assigned in the slot allocation step based on the node data requirements and priority [18].

Mahendran, N., et al., (2020), the researchers suggested a Packet Rank Based Data Scheduling (PRBDS), a new algorithm for optimizing and ranking the incoming data packets depending on multi-attributes decision-making approach, and packet ranking and the schedule was conducted by PRBDS for ensuring data packet delivery within the stipulated time for reaching the base station. Packet priority, size, and deadline are the metrics used by the algorithm for ranking the incoming data packets [19].

Ramadhani Sinda et al., (2020), the researchers suggested an Energy-Efficient Scheduling using the Deep Reinforcement Learning (DRL) (E²S-DRL) algorithm in WSN. E²S-DRL contributes three phases to prolong network lifetime and to reduce network delay that is: the clustering phase, duty-cycling phase, and routing phase. In the first phase, the clustering operation functions via the ZbC scheme that is executed through a hybrid PSO and AP algorithm that reduces energy consumption during data aggregation. The second phase enhances network lifetime through duty cycling. Duty cycling is performed using a DRL algorithm

that schedules each sensor node in a distributed manner. Each scheduling slot considers three modes that are sleep, listen and transmit. In the last phase, routing is performed to reduce data transmission delay by finding the best path between source and cluster head. Ant Colony Optimization (ACO) algorithm is used to choose multiple paths between source and cluster head [20].

The following table (1.1) presents how the scheduling process can be used to conserve energy and extend the lifetime of wireless sensor networks, highlighting the Proposed method, Proposed Algorithm, Evaluation Metrics, and Simulation results, arranged by the publication year.

Table 1.1: Summary of published researches scheduling in WSNs

Ref. No.	Method and algorithm Proposed	Evaluation Metrics	Simulation results
[11]	Data Aggregation → (PFF)	Percentage of lost measures, data accuracy: overall energy dissipation	Results show significantly reduces the number of duplicate data and outperforms in terms of data accuracy and energy consumption.
[12]	Scheduling and, clustering → (EBMS)	Average energy consumption, and network lifetime.	The experiments proved that EBMS outperformed the state-of-the-art approaches.

[13]	Routing, and Scheduling → (ENS_OR)	Average of Residual Energy (ARE), Standard deviation of Residual Energy (SRE), network lifetime, First Dead Node (FDN), and Received Packet Ratio (RPR).	The simulation results show that the new modified ENS_OR with sleep mode conserves the network energy and improved.
[14]	Sleep-scheduling strategy based on Reinforcement Learning → (RL-Sleep)	Packet delivery ratio, average node-energy consumption, sustainability of the network.	Simulation results exhibit a good performance of the proposed approach in terms of sustainable operations of the network and connectivity.
[15]	A Content-based Adaptive and Dynamic Scheduling scheme → (CADS)	Live-Node, status of each sensor's component, number of CH selection procedure.	The results show, it increases energy efficiency in terms of network lifetime by 9.65% in 100 nodes-network, 11.36% in 150 nodes-network, and 0.94% in 300 nodes.
[16]	Scheduling and routing → (JPASR)	Network lifetime, energy consumption, security index.	Extensive simulations results demonstrate that the energy-saving and security performances of the proposed method are superior to the existing ones.

[17]	Scheduling, and clustering → (EASS)	Energy efficiency, network lifetime.	The results show that EASS increases energy efficiency by 50% due to the four-states model, 62.5% due to alive nodes, 92.60% due to minimized Cluster-Heads overhead, and 38.11% due to the reduction in dead nodes ratio.
[18]	Routing, and scheduling	Network lifetime, the residual energy.	The simulation results show that the proposed protocol not only saves energy but also improves channel utilization.
[19]	Packet Rank Based Data Scheduling → (PRBDS)	Energy efficiency, delay packet delivery ratio, lifetime, packet dropping ratio.	A simulation result shows PRBDS not only provides less energy consumption but also significantly reduces the packet drop and increases the lifetime.

[20]	Schedule, clustering, and routing → (E ² S-DRL)	Network lifetime, energy consumption, throughput, delay.	The simulation results reduce energy consumption, reduce delays by up to 40%, and enhance throughput and network lifetime up to 35% compared to author methods.
Our work	Compression, clustering, and scheduling	Active sensor ratio, transmitted data ratio, energy consumption, percentage of lost data.	The simulation results show decreased the transmitted data from 94.3% up to 98.8%, decreased the transmitted data from 75.9% up to 78.7%, decreased the consumed energy from 99.5% up to 99.8%, and decreased the consumed energy from 95.3% up to 98.1% compared to the PFF.

1.6 Thesis Organization

The following parts make up the remainder of this thesis as follows:

Chapter 2 – Wireless sensor networks are described in detail in this chapter.

Chapter 3 - Details a description of the proposed method which is an energy-efficient scheduling-based clustering approach.

Chapter 4 – This chapter analyses and explains the results and also discusses the validity of the results. To evaluate the performance of proposed approaches several performance parameters are used.

Chapter 5 – This chapter presents the conclusion for the thesis as well as some suggestions for future works.

Chapter Two

Theoretical Background

Chapter Two

Theoretical Background

2.1 Introduction

This chapter presents a theoretical background of the WSNs including their architecture, applications, challenges, and energy-saving technologies to save the energy of the sensors and prolong the Network Lifetime. In addition, different types of data correlation that are usually exploited to energy-saving and network lifetime extension are provided. As well as, some of the performance metrics that we used to assess our work.

2.2 Wireless Sensor Networks (WSNs)

WSNs are self-configured networks that are containing a large number of sensor nodes (see [Figure 2.1](#)). These sensor nodes are generally fitted with low-power batteries [\[21\]](#), which are often difficult, costly, and impossible to replace after nodes are installed, furthermore it, sensor nodes can monitor a wide range of physical characteristics such as temperature, radiation, light, noise, humidity, and other characteristics. These nodes gather data and send it one-hop or multi-hop to a Base Station (BS).

In addition, these sensors have a small size, low cost, and wireless data transfer capability. As a result, sensor networks are projected to be widely utilized in a variety of application areas, including for monitoring purposes, in healthcare [\[22\]](#), military applications [\[23\]](#), environment monitoring [\[24\]](#), smart homes [\[25\]](#), smart factories [\[26\]](#), and disaster management. Although WSNs are considered extremely dynamic ad-hoc networks, network topology management, in particular resource

management, scalability, dependability, and efficiency, has been a major issue in these networks [27].

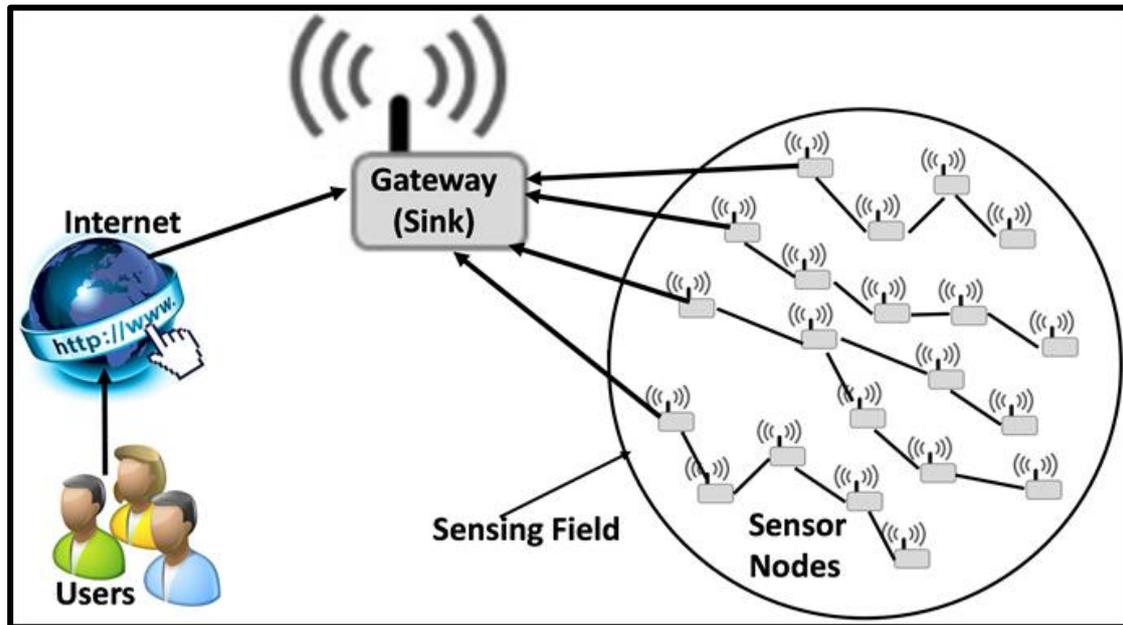


Figure 2.1: Wireless Sensor Network Architecture [28].

1) Node Architecture

The common architecture of WSN's node is described in Figure 2.2. The components of the sensor node are shown one by one. A "Processing Unit", "Sensing Unit", "Power Unit", and "Communications Unit" are the 4 main components of a typical node [29]. The communications unit, which involves data transfer (reception and transmission) [30]. Uses a considerably high proportion of available energy than the other components. The following sections will go over each component in detail.

1. **Sensing units:** responsible for sensing the data from an environment where the node deployed, process it and sends it to the communication unit to get transmitted via the transceiver.
2. **Processing unit:** The sensor node's core is the processing unit. This unit, in conjunction with a small storage unit, controls the processes

that allow the sensor node to interact with other nodes to execute sensing duties.

3. **communication unit:** it is represented by the transceiver, which receives the data from the sensing unit after the processing and transmits it via wireless links. The transceiver has the ability for bidirectional communication, however, some applications need only transmission (TX) or reception (RX) capabilities.
4. **Power Unit:** energy consumption is a basic weakness of sensor networks. methods for conserving energy might in extending the lifetime of the sensors. The sensor nodes are typically powered through either energy storage devices or energy scavenging. The first technique uses a set of tiny batteries to provide power. The latter technique uses energy scavenging from the environment to keep the sensor node running continuously. The most widely used energy provide technique is solar radiation.

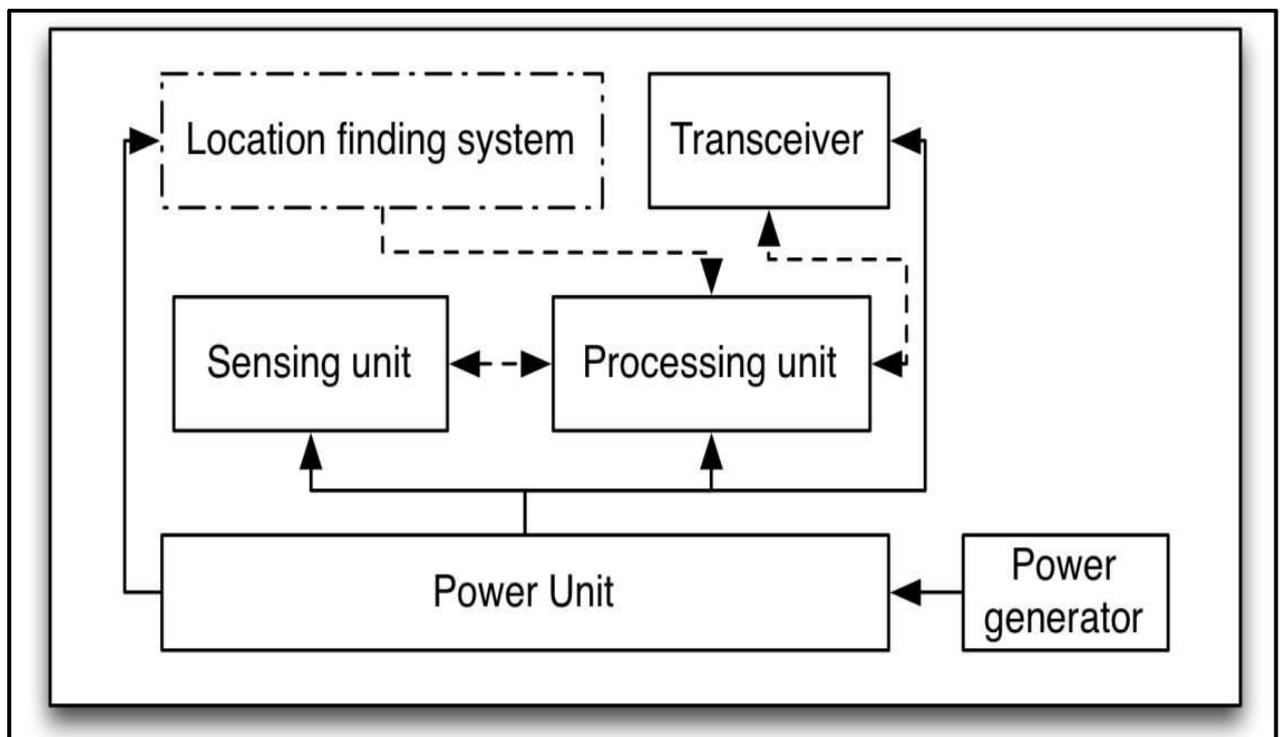


Figure 2.2: The Components of a Sensor Node [31]

Additional components, like a position locating system, a mobilizer, and a power generator, are dependent on the application. A location-finding system is usually included in a sensor node since most sensor network routing strategies and sensing activities require high-accuracy position knowledge. A location-finding system is generally incorporated in a sensor node. A mobilizer could be required in some cases to move sensor nodes whenever needed to complete the assigned function.

2.3 WSN Applications

WSNs are collections of small-size, inexpensive nodes that monitor a specific environment and transmit the data to a Base Station (BS) for more processing. WSNs support a wide domine of important and useful applications thanks to their advantages such as ease of deployment, safe of use, and low cost. Subsequently, an extensive variety of applications are potential as represented in [Figure 2.3](#). and which will be presented briefly in the following section.

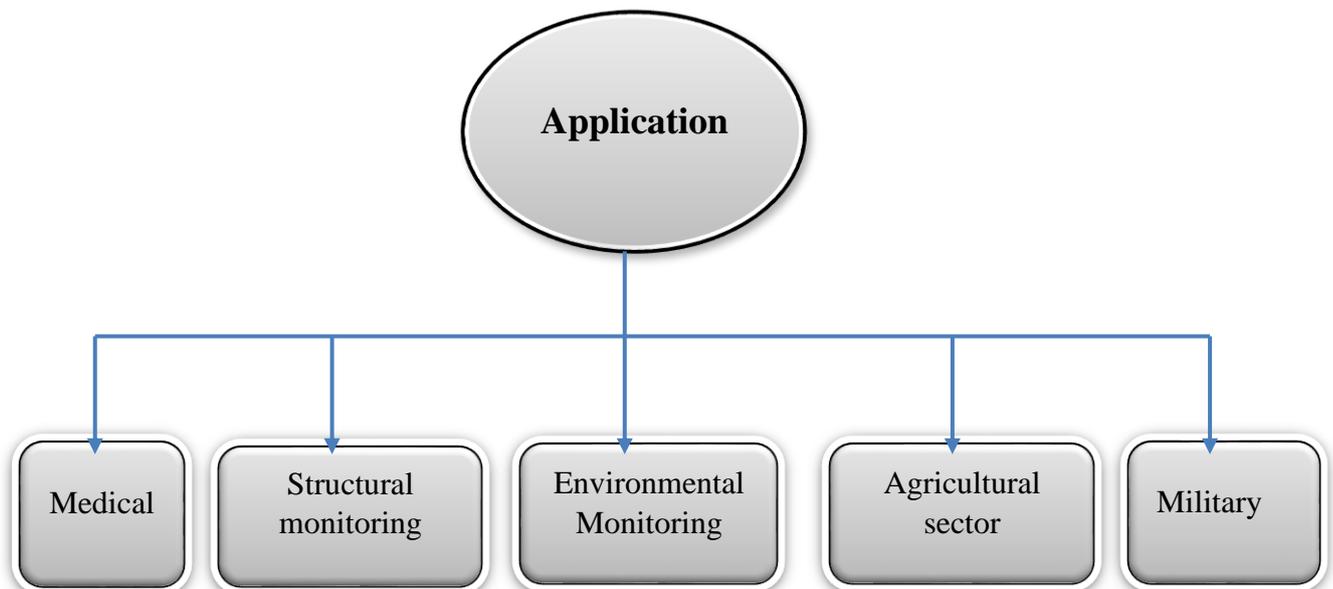


Figure 2.3: Wireless Sensor Network Applications

2.3.1 Medical Applications

Sensor nodes are very useful in diagnosing and monitoring patients. since the patient's physiological parameters, like heart rate and blood pressure, are monitored using sensors implanted in the body. Since the data so collected is sent to inform the doctor of the detected event. this gives patients a big deal of freedom of movement instead of staying in a hospital bed for a long time [32].

2.3.2 Military Applications

WSNs are becoming one of the integral parts of the military command, intelligence, communications, control, and surveillance. Sensors may be utilized to detect, monitor, and track the presence of soldiers and vehicles, allowing for close surveillance of enemy forces [33] [34].

2.3.3 Environmental Monitoring

WSNs have been developed for use in Environmental Monitoring and earth science research. This includes flood detection, sensing volcanoes, gas leakage, Forest fires detection, etc. The data is sensed using the sensors and is sent to the BS via the Internet. As the network provides early detection and prediction of all these environmental events, it helps in taking a safety measure on time and at a high level of accuracy. which helps in taking the necessary measures by specialists before a calamity occurs [35].

2.3.4 Agricultural sector

Sensor networks can help with a variety of activities in the field of agriculture. Used to monitor operations in the agricultural field like measurement of moisture level in soil, watering of the plants, crops

harvesting, pesticide distribution, animal control, and so on. For example, the sensors in a humidity level measuring scenario monitor the humidity level in the crops or soil and wirelessly communicate data to the Gateway site. The necessary actuators can be engaged based on the specified threshold values, such as the water pump motor turning on and off, the motor pump speed control, crop harvesting, crop sowing, pesticide distribution across the crops, and so on [36].

2.3.5 Structural monitoring

Structural monitoring applications can benefit from the usage of sensors. The goal is to develop "smart structures" which can self-diagnose issues and prioritize necessary repairs. This technique is appealing to earthquake-prone areas. Although mild tremors may not cause visible damage, however, it can develop and lead to cracks that can cause the building to collapse in the event of an earthquake of greater strength. so, sensors are useful in this regard.

2.4 Challenges in WSNs

There are many challenges and issues placed on the deployment of sensor networks. That will impact the design, efficiency, and performance of the WSNs, now let bring up briefly some design challenges.

2.4.1 Energy

Sensor nodes require energy for many tasks. Where a lot of energy is expended in collecting, processing, and transmitting data. The sensors are normally powered by a battery that has to be recharged or replaced after or when its power runs out. For certain cases, neither of the solutions are appropriate for modifying or recharging sensor batteries due to cost

constraints or the installation of sensor nodes for extreme, dangerous, or remote regions. Consequently, energy conservation is the most often important problem for a WSN [37, 38].

2.4.2 Security

The deployment of sensors in an unattended area makes the networks exposes to several possible attacks. Therefore, privacy is crucial in WSNs to save the data and information that goes between the network's sensor nodes or even between sensors and the base station; otherwise, communications might be spied on, resulting in a security breach [38, 39].

2.4.3 Quality of Service

The level of service provided by the sensor networks to its users is known as Quality of Service (QoS). The authors describe QoS for sensor networks as the optimum number of sensors sending data to sinks or a BS. The QoS in WSN is difficult because the network topology may constantly change. This dynamic nature of sensor networks makes the availability of precise state information next to impossible [39].

2.4.4 The Dynamic Topologies

In industrial environments, because of link failure or node location change, the topology and connectivity of the networks may change. In addition, sensor nodes may also be exposed to high humidity levels, vibrations, dust, and dirt, as well as other circumstances that impair performance. These external factors and the dynamic topology of the network may cause certain sensor nodes to fail [40].

2.4.5 Hardware Limitations

At the absolute least, each sensor node must have a "Sensing Unit", a "Processing Unit", a "Transmission Unit", and a "Power Supply". the

nodes could contain several built-in sensors or external devices like a system of localization. Every additional function, however, comes at a cost, increasing the power consumption of the node. Therefore, added functionality has to be weighed continually against costs and power consumption [41].

2.5 Energy-Efficient Mechanisms in WSNs

WSNs are employed in many applications because sensor nodes in such networks have limited battery life, therefore, energy becomes the most precious resource for them [42]. To optimize the usage of energy resources, thus the research community is attempting to provide energy-efficient solutions for WSNs to extend the network's lifetime. The following are some of the most well-known energy-saving techniques:

2.5.1 Schedule Mechanism

As previously stated, sensors are energy-constrained, and conserving their energy is critical as the lifetime of the network is limited by battery lifetime. In many cases, recharging or replacing the sensor node's battery is not possible. In addition, sensors are usually deployed in huge numbers, so they don't need to work with all sensors at the same time. Working schedules including active and sleeping schedules can be considered to run the sensors more effectively since the sensors consume much smaller energy in their sleeping time compared to the time, they are being active. In such a case, rather than having all sensors transmit the similar data, it has a higher efficiency for the selection of a few representative nodes to tell the sink node about an observed event.

Sensor networks have different working modes, i.e., (1) active mode, (2) idle mode, and (3) sleep mode. In active mode, the sensor can "Collect Data", "Send/Receive messages", "Process Data", and do other

computations [43]. When the node is in idle mode, it continues to listen to the medium even when no messages are being sent and it consumes tens to thousands of times the current when compared to sleep mode. When the sensor's CPU is switched off, this mode is also known as the Sleep mode.

For that, the motivator for establishing sleep scheduling mechanisms is to minimize the idle state energy consumption, focusing on sleep mode that will prolong the network lifetime. Several methods are proposed to address the challenges of idle listening, for example.

A. Duty Circling

When designing WSN consisting of nodes with limited resources and battery-powered, we expect maintenance-free operation for long years. Therefore, the node must temporarily turn off its radio to reduce energy consumption and ensure adequate lifespan. The technique in which the nodes go periodically into sleep and wake up only for a short period is called duty cycling [44].

Duty circling, also known as the “sleeping schedule mechanism” is the part of the time in which the sensor is in the “sleep” working state in a pre-defined period. The main purpose of duty circling is to minimize the idle mode that is the dominant factor in most sensor network applications. A robust and suitable duty circling mechanism can prolong the lifetime of a WSN [44]. Carrano et al.[45] have created a survey of duty-cycling mechanisms, dividing them into synchronous, asynchronous, and semi-synchronous as shown in Figure 2.4 [46].

- **Synchronous Wake-up Mechanism:** In this type, all the sensors use a common time reference i.e. all sensors turn on or off in the same period to exchange information. Although this kind of is easy and uncomplicated, it faces

many problems due to the synchronization delay as well as extra control traffic [44].

- **Asynchronous Wake-up Methods:** Asynchronous sleep-wake scheduling refers to the fact that nodes wake up to transmit regardless of whether or not other nodes are awake to receive. Because of this, Asynchronous schemes do not need time Synchronization [47]. it is an effective way to save energy by appropriately arranging sensor nodes to sleep or wake.
- **Semi-synchronous Wake-up Method:** This type is also called hybrid, where the switches between the two types mentioned previously, depending on the conditions of the network. The choice to switch between the two modes is based on the idea that asynchronous methods perform better under light load situations, while synchronous methods perform better under heavy load conditions [47].

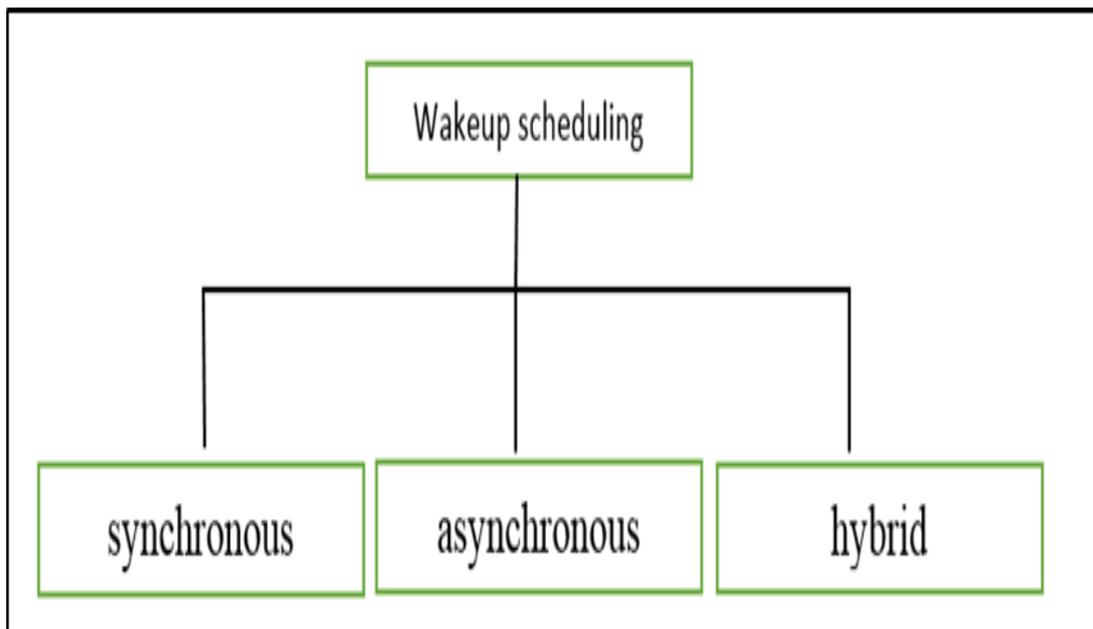


Figure 2.4: Wake up Scheduling Schemes classification

B. Active and Passive Wake-Up Radio

Currently, used radio transmitters consume approximately the same amount of energy when they receive data and when they listen to the channel waiting for upcoming transmissions (idle listening). Hence, to reduce energy consumption the WSN nodes switch the radio between active and sleep modes. In the active mode, a node can receive and transmit packets, while in the sleep mode it completely turns off its radio to save energy. Since frames cannot be exchanged unless both the transmitter and the receiver are active, the node must either be aware of its neighbors' wake-up time or be able to wake up its neighbor node. The common technique used in WSN is duty-cycling when the node sleeps most of the time and wakes up periodically only at scheduled moments. These moments are known to the transmitter which must wait a certain time for waking up its neighbor to exchange data. This issue has a direct and negative impact on the latency of the network. The scheduled duty-cycling technique is always a trade-off between energy consumption and delay in delivering the frame to the target node. The delay problem can be alleviated by an additional wake-up radio (WuR) [48]. The idea is to use an additional receiver with such a low power consumption that it can be active all the time. A node will be woken up only on demand, which decreases communication delay and which is important in time-critical applications [44].

2.5.2 Clustering

Clustering methods for WSNs have been proposed as a major energy conservation goal in the last years [49]. Data clustering can be defined as a "Machine Learning" approach that groups a set of data points into clusters such that points within each cluster are similar to each other [50]. it is a

common method that divides a network into two layers. The nodes located in the same layer are grouped into clusters. To efficiently distribute the management tasks among the nodes, some of them are elected to be the head of each group (cluster), which are usually called cluster heads (CHs). Because large amounts of the data created by sensors in WSN are comparable, clustering takes advantage of data correlation and aggregates it to reduce network load, leading to more effective energy consumption.

The common architecture of clustering is shown in (Figure 2.5.) Here, clusters are provided with "Cluster Heads", and the sensors in the first layer sense the region and then transmit the data to their related Cluster Head. The CHs in turn, receive this data in the 2nd layer, and after carrying out several operations, and send it to BS. Eventually, the user receives data from BS via the Internet. In this way, we will reduce the power needed to transmit and receive, and balance the power consumption.

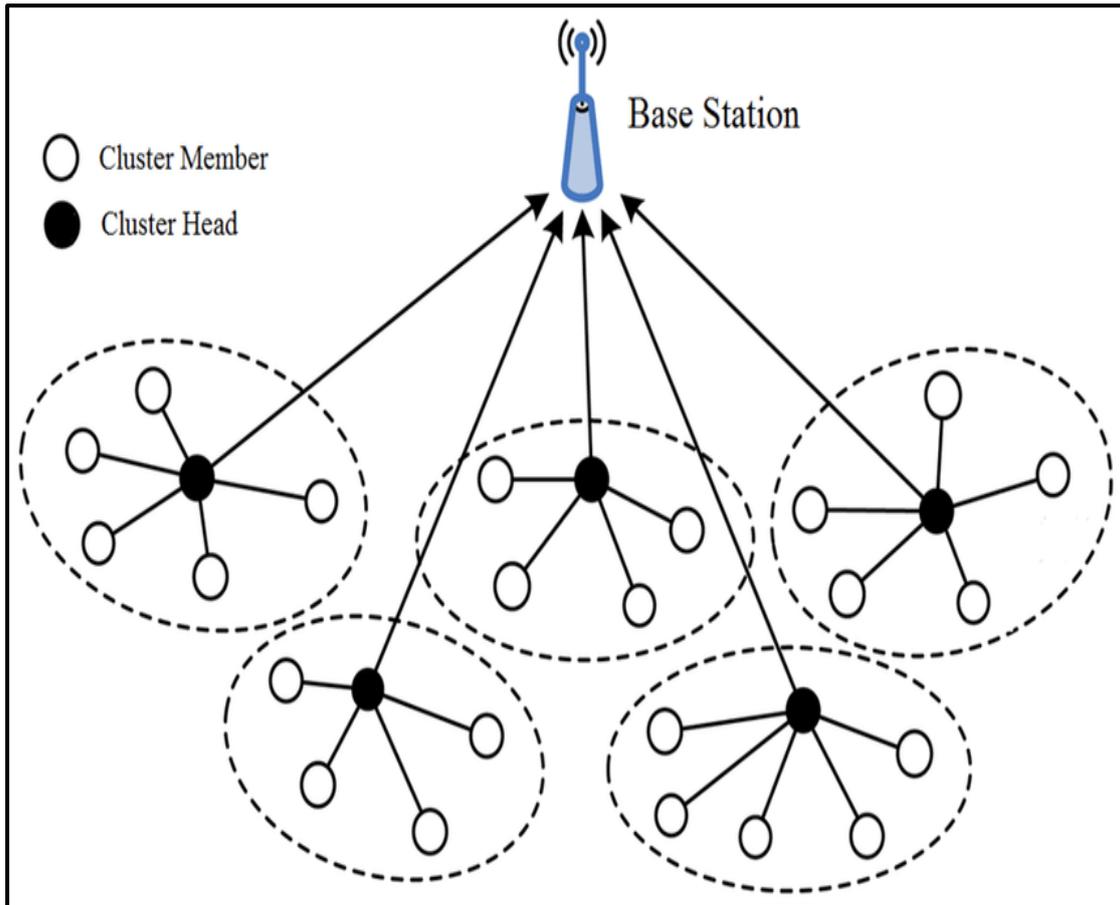


Figure 2.5: Clustering in Wireless Sensor Networks [51]

With clustering approaches, two phases are often taken into consideration: **CH selection**, and **cluster formation**. The important aspect in any approach is the CH selection process which determines the network's energy efficiency. In addition to the energy efficiency, other objectives such as full-field coverage might be pursued as the main goal of the clustering strategy in addition to energy-saving [49].

The following section will be presented general CH selection methods, objectives of clustering, and clustering algorithms.

A. General CH Selection Methods

Cluster Heads (CHs) selection strategies might enhance the network efficiency by identifying nodes that are better suited to be CHs [52]. There are several general ways for selecting CHs as in the following:

- Resource-rich nodes are predefined as CHs in several clustering approaches [53]. The issue with this approach is that the majority of WSNs are resource-constrained and homogeneous. As a result, the approach is inapplicable in several situations. Furthermore, even in the case when a resource-rich node in a heterogeneous network might be identified and designated as a cluster head, being a CH for an extended period can rapidly deplete the node's power, resulting in node death.
- In many clustering approaches, randomization is the solution for choosing who will be CH among nodes. While such technicality is important in inhomogeneous networks, any unbalancing or dynamicity in WSNs might result in severe problems, including runtime issues such as unbalanced resource consumption or continuous energy consumption in a few CHs.
- The most common approach for cluster head selection has been referred to as the conscious CH selection, and it is mostly dependent on network and nodes conditions. This concept has been proposed in many articles in literature as a basis for clustering approaches. CHs are chosen using such procedures depending on a variety of factors like the location, available resources, number of neighbors, and so on.

B. Objectives of Clustering

Clustering the nodes in WSNs is performed with different purposes and objectives. The most significant and common goal of all such goals is energy conservation. These objectives are classified into primary and secondary. The primary objectives are the most important and substantial in the clustering process. On contrary, the secondary objectives are not substantially important for the network and they have been indirectly achieved through clustering the nodes [54].

A few of such goals are briefly described as follows:

- 1. Maximize network lifetime:** As mentioned previously, extending the lifetime of the network as much as feasible is a major difficulty in WSNs. A clustering process that meets such goals can help extend the network's lifespan. For instance, in the case when CHs are positioned in the node population's center, the role of CH is correctly rotated between nodes, and sleeping schemes are efficiently exploited, then we can expect a good life for the network.
- 2. Data aggregation/fusion:** This is an important technique to prevent the transmission of repeating data in the network since a large amount of data in WSNs is the same.
- 3. Scalability:** The scalability of networks can be increased well by dividing them into some clusters. In the case where a node in the cluster requires to connect with a different node in a different cluster, it has to know certain information regarding the related cluster head of the cluster where the other node has been located. Which improves the network's scalability and results in a large reduction in routing table sizes.
- 4. Stabilized network topology:** Because the nodes are divided for some of the clusters, CHs can control changes in node position at the

cluster level. Each CH in a clustered WSN has information regarding its members, like their location and energy; thus, if a node moves to another cluster or dies, the CHs are notified right away.

- 5. Increased connectivity:** This goal might be as simple as possible, guaranteeing that a path exists from each cluster head to base station, or it might be more restrictive through setting a bound on path's length. In the Worst-Case situation, the k -connected network will be disconnected if k nodes fail. Particularly in large-scale WSNs, clustering the nodes enhances performance. This is because, in contrast to a flat architecture with clustered nodes, there must be a minimum of one path from each cluster head to the base station for achieving connectivity, not between every node in the network and base station.

C. Clustering Algorithms

Clustering algorithms are divided into 2 types, which are: equal-size clustering and unequal-size clustering.

- **Equal-sized clustering algorithms:** Equal-sized clustering approaches have extensively been explored by many researchers. The main idea behind this type is creating clusters that are of rather equal sizes, to keep the number of the clusters to the minimum, distributing them evenly over the network.
- **Unequal-sized clustering algorithms:** Typically, in unequal clustering, there is a hot-spot problem in which Cluster Heads closer to BS tend to die faster, the basic reason is that cluster heads that are located in BS's vicinity have to relay more data in comparison with the farther cluster heads. As a result, energy is dropped at a faster rate. Therefore, these algorithms are suggested for balancing

consumption of the energy amongst the clusters, where sizes of the cluster near sink node are considerably smaller in comparison with clusters that are far from the sink node to save more energy in the intra-cluster and inter-cluster communications.

2.5.3 Energy-efficient Routing

Sensor nodes have limited energy and bandwidth. To optimize the lifespan of sensor nodes, the network layer tries to find ways to establish up Energy-Efficient routes and reliable data relaying from sensors to sink. This provides a balanced power consumption for the networks. Various data collection algorithms are used to reduce transmitting distances for such networks through routing [55]. An example of some data routing methods is demonstrated below [56].

- 1. Hierarchical Routing:** In this method, sensors forward their readings to higher-level nodes to be sent finally to the base station (BS). Clustering and tree-based routing are included in this classification. Sensing data is transmitted in multiple hops to deal with the long distance from all sensors to the BS.
- 2. Flat Routing:** in this type, all the nodes in the network are given equal roles to perform the sensing tasks. For example, random walks (RWs) are randomly chosen among sensors to initiate some routing. Sensors randomly choose one of their neighbors to forward their readings for the next RWs. Through RWs, the sensing data will be collected at the BS until the last RWs reach the BS, also randomly.
- 3. Location-Based Routing:** Before sending sensor readings to the BS, several data collection approaches for the WSNs require location information for the sensors for the calculation of the data

transmission distances between 2 sensors or across the entire network. According to this, the amount of energy used for the transmission of the data may be calculated and chosen, allowing for the lowest energy consumption. Sensors' local information can be used to route data in an energy-efficient manner.

2.5.4 Energy Harvesting Techniques

The main energy source of the WSNs is mainly based on battery power. When the battery runs out, the sensor cannot complete its task [57]. Besides, may lead to early death to the network. To overcome these problems and also extend the operating time for the sensor node, the energy from the surrounding environment is utilized .

Energy harvesting (**EH**) is the process to obtain energy from the outside environment like the wind, vibration, flow energy, acoustic, and thermal energy, solar, and more. The technologies utilized in energy harvesting convert the energy from the environment to electrical energy, which may be utilized in the WSNs [58]. EH system opens up potential technologies for the WSN environments that remotes the dependence of sensor buttons on battery power, significantly improving the cost-effectiveness and performance needed for the WSNs. The combination between the EH source and sensor node architecture is depicted in (Figure 2.6) EH sources are becoming the current way, especially for small and low power devices such as WSNs to be a complete replacement for battery use.

On the other hand, highlight the continual unavailability of energy harvesting sources; as a result, a buffer has been suggested to store energy for later uses, employing a battery-free sensor node and low-powered capacitors as the buffers. Solar energy, for example, is unable to gather at

night, because of the absence of the sun. This presents a problem. As a result, special devices of energy harvesting could be put in-network for scavenging the energy and transferring it to nodes wirelessly.

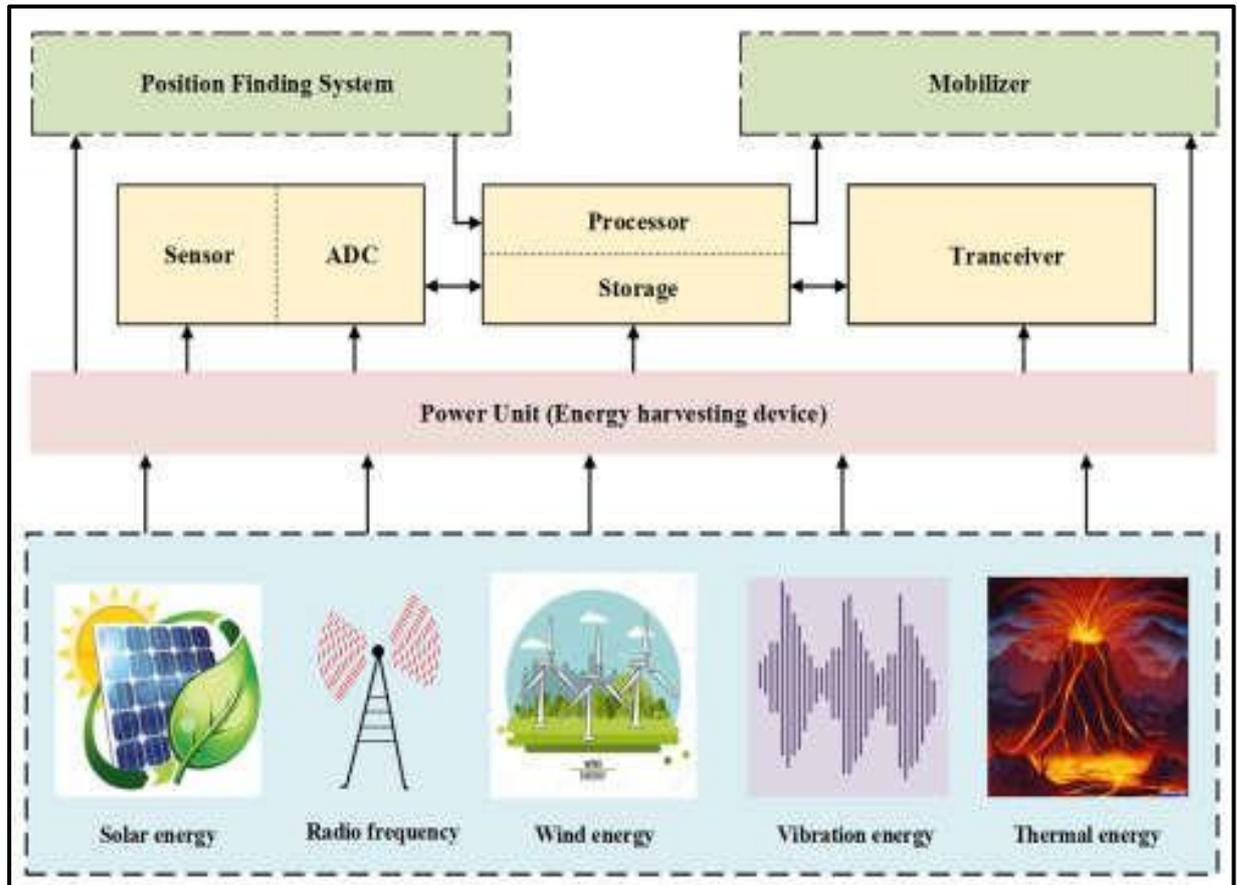


Figure 2.6: Energy Harvesting Techniques [56]

2.5.5 Data compression

Because data communication has been considered as the most energy-exhausting task which is undertaken by a sensor, and sensor nodes have limited energy, sending data directly to the sink is wasteful; additionally, reducing energy consumption in each sensor node has become necessary to improve the lifetime of WSN. Consequently, we want strategies for removing redundancy and consolidating data to produce information with

high quality at the sensor nodes which will minimize the number of packets that will be sent to the sink and thus lead to conserving both energy and bandwidth [59]. To achieve this data compression can be used.

Can be defined Data Compression as the process of coding data to reduce its size [60]. Data Compression is one of the most important Energy-Saving processing methods it will decrease both: the transmissions and receiving energy.

The studies previous and current illustrate that the communication process is has a higher power consumption compared to the calculation, Sensing, and Processing, therefore, minimizing the size of the data before the transmission in wireless medium has been found efficient in the reduction of the total consumption of the power. This is why it is advantageous for the WSNs to utilize approaches of data Encoding.

"Data-Encoding" techniques hold the promise of decreasing communication energy [61]. This is done by compressing the sensor data before sending it without affecting the integrity of the data. Due to the limitations of computing resources such as RAM, the algorithm also needs to be lightweight and computationally efficient. In this thesis, we use two algorithms for encoding (as described in the practical part). One of the ways encoding data is Run Length Encoding (**RLE**), and the other is Huffman. and will explain this below:

Data compression is commonly utilized in a wide range of areas. The fundamental data compression goal is minimizing the number of bits that are needed for coding information or methods of data compression is categorized generally into 2 classes, which are: Lossy and Lossless compression.

- **Lossy compression:** those methods encounter information loss while decompressing compressed information. Those approaches may be implemented in the images, graphics, and digitalized audio. Many of those are altered to various quality levels, giving higher precision in transferring fewer effective compressions [62].
- **Lossless compression:** Those methods accurately code data and transfer them; there won't be any loss type in the data whereas decompressing compressed information [62]. Which is applied for storing database records, word files, spreadsheets, and so on. Huffman coding and RLE are examples of approaches utilized in the compression of data [63].

This section reviews only lossless compression techniques used in this thesis, such as **Huffman** coding and **RLE**.

A. Encoding RLE

Run-Length Encoding: As has been stated earlier, sensor nodes may take the same (or very similar) readings several times when a time slot is very short. This is why results will include a repeating string of similar values. This is why, to reduce the amount of data transmitted, the RLE encoding approach is employed for more compression. RLE is a very basic and powerful encoding strategy. The nearby grouping of symbols that happen in succession is exploited by RLE. and replaces" run" of symbols by a tuple containing the symbol and the number of times it has been repeated. For instance, considering the string, AAABBBBBBCCCC is encoded as [A,3] [B,6] [C,4], and will explain the algorithm used in chapter 3.

B. Huffman coding

Huffman coding has been considered the most common approach for the generation of prefix-free codes. Where this coding is commonly utilized in data compression areas. This is why the Huffman algorithm is one of the optimum lossless data compression methods. In Huffman coding, symbols that have a high likelihood of occurrence have been more emphasized. This compression approach has been found sufficient for the data for which there is a high occurrence frequency. Where Huffman's algorithm [64] works from the leaves to the root in opposite orientation [63]. Algorithm 2.1 introduces the Huffman encoding algorithm implement in sensors level. and Algorithm 2.2 introduces the Huffman decoding implement in gateway level.

Algorithm 2.1: Huffman Encoding

Input: D: input file, Mx

Output: CF: compressed file.

```
1: LT ← QUEUE ();
2: for j ← 1 to Mx do
3:   Nd ← Node {Dj};
4:   LT.PUSH(Nd);
5: end
6: while LT.Size () ≠ 1 do
7:   Nd ← NewNode ();
8:   Nd.Left ← X ← LT.POP ();
9:   Nd.Right ← Y ← LT.POP ();
10:  Freq (Nd) ← Freq (X) + Freq (Y);
11:  LT.PUSH(Nd);
12: end
13: CF ← LT
14: return CF;
```

Algorithm 2.2: Huffman Decoding

```
Input:  Rt: the root of Huffman tree,  
         Bt: bit stream is needed to be decoded.  
Output: Df: decompressed file.  
1: Ms ← Sizeof (Bt);  
2: for j← 1 to Ms do  
3:   Df←Rt;  
4:   while Df. LEFT≠ NULL and Df. RIGHT ≠ NULL do  
5:     if (Btj = 0) then  
6:       Df←Df.LEFT;  
7:     end  
8:     else  
9:       Df←Df.RIGHT;  
10:    end  
11:    j←j+1;  
12:  end  
13: end  
14: Return Df;
```

2.6 Data Clustering

Clustering algorithms are widely classified into two types in recent research: hierarchical and partitional [65].

2.6.1 Partitional Clustering Algorithms

Those partitional approaches have been common in a variety of study areas as a result of their ability for clustering large datasets in WSNs for the classification of sensors to improve coverage and lifetime. In addition, The K-means algorithm can be defined as the most basic partitioning concept of clustering known yet, it remains commonly utilized and popular for high dimensional datasets as a result of its simplicity and

the lower computation complexity [66, 67] more about it will be explained later.

k-means: The k-means clustering approach is an unsupervised algorithm of machine learning, it is the most widely used partitional clustering algorithm. It is utilized for the identification of the clusters of the data objects in the dataset. It begins with the selection of K representative points as initial centroid values. Every one of those points is assigned after that to the nearest one of the centroids, according to a specific measure of proximity that has been selected. As soon as the clusters have been formed, the centroids for every one of the clusters will be updated. After that, this algorithm performs an iterative repetition of those 2 steps to the point where centroids don't change or another alternative relaxed criterion of convergence is met [68].

2.6.2 Hierarchical Clustering Algorithms

The hierarchical algorithms of clustering have been developed for overcoming some drawbacks that are related to the partitional-based approach of clustering.

The hierarchical approaches may classify as "Agglomerative Hierarchical Clustering" (**AHC**) and "Divisive Hierarchical Clustering" (**DHC**) approaches. The agglomerative method starts by taking the singleton clusters (containing a single object in each cluster) at the bottom level and keep to merging 2 clusters each time to build a bottom-up hierarchy of clusters, more precisely it merges the two points that are the most similar until all points have been merged into a single cluster. It is also called the bottom-up approach. Divisive approaches, however, begin with all data objects in a massive macro-cluster and continuously split them

into 2 groups, generating a top-down cluster hierarchy, it is also called the top-down approach. Agglomerative methods are the focus of our attention in this thesis and it will be explained in more detail later.

AHC is one of the important traditional unsupervised hierarchical clustering methods, it is a bottom-up method that starts with clusters, each of which includes exactly one object. It recursively partitions the dataset into a tree structure through a series of merge operations based on proximity measures. And finally, it forces all the clusters into a single cluster [69].

2.7 Data Correlation

Data collection is an important application of WSNs. For example, in environmental monitoring, the physical variables, such as temperature, light illumination, and humidity, etc. are often sensed and sent from sensor nodes to the sink via "Single-Hop" or "Multi-Hop" routing. Because sensor nodes typically have limited power supplies, one of the major goals of data collection is to get the needed sensing data with the least amount of energy consumption, therefore extending the network lifetime. Since the physical variables describing natural occurrences are usually locally regular and WSNs nodes are often deployed redundantly to get complete environmental monitoring data, the sensing data often have a high spatial or temporal correlation. As a result, several researchers employ spatial and temporal correlation to reduce network energy usage by sending only a portion of sensor data or aggregating it [70, 71].

- 1. Spatial correlation:** WSN applications require spatially dense sensors deployment for achieving required coverage. As a result, several sensors in the sensor field collect similar data about a single event. In this case, instead of having all sensors reporting the same

data, it is more efficient to choose a few representative nodes to report the sink node about the detected event. A representative node reports the event information of a given specific region on behalf of a collection of nodes that gather similar information in the same area.

2. Temporal correlation: many applications of the WSNs may need a lot of sensors to periodically perform the observation. As a result of the physical event's nature, there's an important temporal association amongst every consecutive sensor node observation. For instance, in daily sampling for the temperature carried out at every minute, the temperature might not significantly vary. In which case, it is not required to report new reading each minute, because the earlier reported reading is matching the actual one.

3. Spatio –Temporal Correlation: When the data collected include both the two previously mentioned correlations. The spatial correlation, due to their proximity to each other, and the temporal correlation because the current reading correlated with the previous one, this is called Spatio-temporal correlation.

Some scheduling solutions can exploit both correlations for node scheduling to decrease the send data to the sink and conserve the network energy.

2.8 Network Lifetime

As previously said, a WSN made up of low-cost, and energy-constrained sensors are used to monitor physical events. As a result, network lifetime becomes an important consideration in WSN design. Several definitions of network lifetime age can be provided as in the below:

1. The term "network lifetime" refers to the time between the moment of deployment and the time when the network is considered nonfunctional.

For instance, when the first sensor dies, or several sensors die, leading to a loss of coverage [72, 73].

2. The lifetime of a network is defined as the time that it takes from the start of the network operation to the end of the energy of the first node within the network [72].
3. It is defined as the maximum period after which a network cannot do tasks as required [74].
4. it has been defined as the time between the first transmissions of nodes within the network into a low percentage of the node compared to the value of the threshold that was determined depending on the type of application [74, 75].
5. Network lifetime based on application requirements some authors consider that network is alive as long as application functionalities are ensured [73].

2.9 Energy Consumption Model

WSNs often consist of large number of linked sensors. It can sense and compute in addition to communicating with each other. The WSNs are systems with significant energy consumption restrictions, and extending the battery life of sensor nodes is a critical need for network autonomy. To achieve this aim, you must first have a deeper knowledge of how energy is consumed in a sensor node.

The methods proposed in this thesis are based on reducing costs (in terms of energy) by distinguishing and reducing redundant data while maintaining the integrity and accuracy of information. The proposed technique employed the energy consumption model proposed in [76]. This model took into account only the consumed energy of transmitting and

receiving. Hence, to send an h -bits message for distance, the model computes the energy consumption as follows:

$$ETX(h, dis) = E_{elec} * h + \beta_{amp} * h * dis^2 \quad \dots\dots\dots (2.1)$$

The consumed energy is computed for the reception as follows

$$ERX(h) = E_{elec} * h \quad \dots\dots\dots (2.2)$$

Where the consumed energy of the electronics of the radio is indicated by **Eelec**. The amplifier consumed energy is described by the β_{amp} . Each sensed data is considered 64 bits. Therefore, the size of the data packet is the number of transmitted readings multiplied by 64 bits.

2.10 Performance metrics

The suggested approach that is explained in detail in chapter three has been evaluated using a variety of performance measures such as Active Sensor Ratio (**ASR**), Energy Consumption (**EC**) in Joules, Transmitted Data (**TD**) in KB, and Percentage of Lost Data (**PLD**). These measures are briefly explained below:

1. **Active Sensor Ratio (ASR):** One key performance metric is to reduce the number of active sensors that take on the sensing task during the next period while maintaining an acceptable data accuracy at the sink node, it is the average number of active nodes in each period.
2. **Energy Consumption (EC) in Joules:** This is a measure of the consumed energy during the transmission and reception in each

period over the network based on the energy consumption model proposed by Heintzelman. [76].

- 3. Transmitted Data (TD) in KB:** This is the rate of data transmitted to the BS by the active nodes.
- 4. Percentage of Lost Data (PLD):** This measure refers to the amount of data that is not sent to the BS (sleep nodes' data) and also, they do not exist in the data received by the sink node from the active sensor nodes. so, we must balance the energy expenditure on the one hand and the percentage of data loss so that there is no significant data loss, which leads to a loss of coverage and the required accuracy.

2.11 Summary

In this chapter, an overview of the WSN has been given. Several application fields such as environment monitoring, military applications, and medical monitoring have been presented also the main challenges in WSN have been described, the explanation of the clustering approach has been illustrated, which is aimed to decrease the consumed energy of the WSN, The Energy-efficient mechanisms in WSNs are shown. Data encoding and Data Correlation are explained. Energy Consumption Model is given. Some definitions of the Network lifetime are provided. Finally, performance metrics are discussed.

Chapter Three

The Proposed Approach

Chapter Three

The Proposed Approach

3.1 Introduction

This chapter provides a detailed description of the proposed Energy-efficient Sensor Activity Scheduling approach for the lifetime enhancement of WSN. The proposed approach work into rounds. Each round is composed of two periods. The first period consists of three steps: data collection, compressed data transmission, and scheduling-based clustering. The second period consists of only collection and compressed data transmission (See [Figure 3.1](#)). The proposed approach is implemented in two levels: sensor nodes and the gateway. The sensor nodes capture the data and compress them using either Modified Run Length Encoding (**MRLE**) or Huffman Encoding (**HE**). At the gateway level, the approach applied the agglomerative clustering approach or k-means algorithm to produce the best schedule of sensor nodes that remain active in the next period. Sensor nodes are clustered based on the received vectors of data into groups and one sensor will be elected in each group to stay active in the next period (the sensor with the highest energy is chosen to remain active). In the following sections, a detailed description of the proposed methods is given.

3.2 The Framework of the Proposed Work

In this section, we will present the proposed approach for extending the lifetime of sensor nodes in WSNs networks in more detail. The proposed approach is described in [Figure 3.1](#).

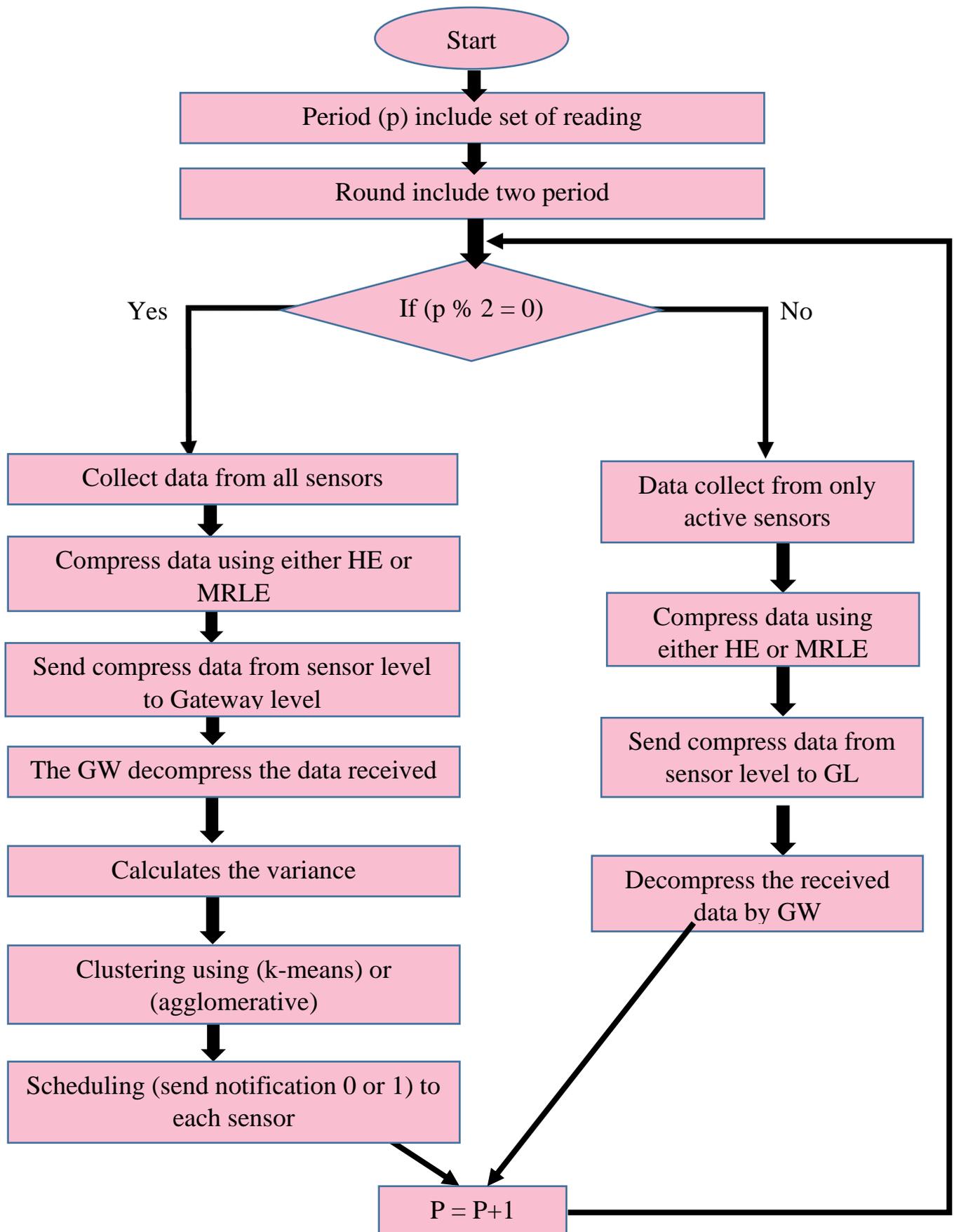


Figure 3.1: Flowchart of Proposed Approach

The goal of this approach is to reduce transmitted data, activate the least number of sensors, and minimize consumption of energy thus extending network lifetime while maintaining suitable accuracy. The proposed approach combines two energy-efficient techniques i.e. data compression and sensor activity schedule which will save energy and reduce the redundant data transferred from wireless sensor nodes to the gateway node.

3.2.1 Sensor level

At this level, the sensor achieves two main functions.

- **Data collection:** In this step, the all sensor nodes are wake up in the first period of each round, sensing the area of interest for a predefined fixed time of the period, and then save these data temporarily for the next step in this period. The data collection in the next period in each round will achieve by only the active sensor nodes that are notified by the gateway.
- **Compressed data transmission:** The objective of this phase is to minimize the amount of collected data before sending it to the gateway by eliminating the redundant data without affecting the data integrity. The collected data is compressed using either Huffman coding (See Algorithm 2.1) or modified Run Length Encoding (See Algorithm 3.1) and then send the compressed file to the gateway. When the gateway gets all of the sensor nodes' compressed data files, it reconstructs these data using the Huffman Decoding algorithm (See Algorithm 2.2). or using modified Run Length Decoding algorithm (Algorithm 3.2) This algorithm is executed for each received file of sensor nodes to return the original data reading. The original LRE is a lossless encoding method deal

with the character, while the Modified RLE is lossy encoding method and the ratio of data loss depend on a predefined threshold. Moreover, MRLE deals with numerical values. [Algorithm 3.1](#) introduces the MRLE algorithm.

Algorithm 3.1 MRLE

Require: **D:** data readings, **th:** threshold

Ensure: **C:** compressed sensor data readings.

```

1: amount ← 1
2: C ← ∅ // storing the encoded string
3: index ← 0
4: while index ≤ (length(D)-1) do
5:   if index = (length(D)-1) or Dif(Dindex, Dindex+1, th) ≠ 1 then
6:     C ← C ∪ Dindex
7:     amount ← 1
8:   else
9:     amount ← amount + 1
10:  index ← index + 1
11:  endif
12: endwhile
13: return C

```

The function Dif return 1 if the difference between the D_{index} and D_{index+1} is less than or equal to threshold th, and it returns 0 otherwise.

3.2.2 Gateway level

In this level, the sensor activity scheduling is applied only in the first period of each round. The proposed approach decodes each compressed data file received from the sensor node k and saved in V^k . It reconstructs these data using the RL Decoding (RLD) algorithm ([see](#)

Algorithm 3.2). This algorithm is executed for each received file of sensor nodes to return the original data reading. After that, the variance D^k of each constructed vector will be calculated, according to variance equation 3.1.

Algorithm 3.2 RLD

Require: C: compressed sensor data readings

Ensure: D: sensor data readings.

```

1: M ← read file C
2: for each row in M do
3:   L = row1
4:   for x in range(L) do
5:     D ← Append (row0)
6:   end for
7: end for
8: return D

```

The row₁ is the number of occurrences of reading row₀ and the row₀ is the reading in the compressed M file.

$$D^k \leftarrow \text{Variance} (V^k) \quad \dots\dots\dots (3.1)$$

Where k represents the sensor node, V^k decompressed data file, and D^k represents a value of variance. The variance is calculated using the equation below (3.2).

$$\text{Variance} = \sigma^2 = \frac{\sum(x_i - \mu)^2}{n} \quad \dots\dots\dots (3.2)$$

Where x_i represents the mean, μ represents each value in the data, and n means the number of values in the data set.

Hence, each sensor node dataset is represented by its variance value. To generate the sensor nodes' schedule that takes a mission of monitoring for the next period, high spatial correlation of the data between sensor nodes reading in WSN is exploited in particular in dense sensor networks, where nearby sensor nodes are supposed to sense similar data. Therefore, the most similar sensor data are clustered in the same cluster.

The role of scheduling here is to keep only one sensor (the sensor with the highest energy) transmitting the data, and the others of the sensors go into sleep mode. The scheduling process here is by simulation process, sending a notification (0 or 1), representing node states (i.e., **active/sleep**) to each sensor node for the phase of the sensing in the current period. so, 0 refers to sleep mode in the next period, 1 refers to active mode in the next period. However, the scheduling process is achieved by k-means clustering algorithm or hierarchical clustering algorithm (agglomerative Algorithm).

A. Scheduling based agglomerative Algorithm

The hierarchical clustering algorithm (As shown in [Algorithm 3.3](#)) is utilized to cluster the close or similar data into groups defined as clusters. The produced clusters are different and the properties of the whole objects which belong to the same group are close or alike. There are two types of this type of clustering: divisive and agglomerative clustering. More details about hierarchical clustering are explained in detail in chapter two ([See 2.6.2](#)). In this thesis, we focused on the agglomerative approach. That it begins from the bottom up by considering each data point as a single cluster. After that, the similarity between every two clusters is calculated. Every two clusters with minimum distance are merged in one cluster in the upper level of the hierarchy. This operation is continued on the upper level of the tree by calculating the distance between every two clusters and merging the ones with minimum distances until providing one cluster. The

similarity method that computes the distance between every two clusters affects the final results. Various metrics can be used for calculating the similarity between every two clusters such as maximum, Manhattan, square Euclidean, and Euclidean distances. It is important to select the metric that is used to merge two close clusters after computing the distance between them. These metrics are single, complete, average, and ward linkages. In the present approach, the ward linkage (According to equation 3.2) has been used by calculating the distance from the differences of the sum of squared of the whole clusters.

$$\Delta (CR_1, CR_2) = \sum_{j \in CR_1 \cup CR_2} \|\vec{y}_j - \vec{m}_{CR_1 \cup CR_2}\|^2 - \sum_{j \in CR_1} \|\vec{y}_j - \vec{m}_{CR_1}\|^2 - \sum_{j \in CR_2} \|\vec{y}_j - \vec{m}_{CR_2}\|^2 = \frac{n_{CR_1} * n_{CR_2}}{n_{CR_1} + n_{CR_2}} \|\vec{m}_{CR_1} - \vec{m}_{CR_2}\|^2 \quad \dots \quad (3.2)$$

Where Δ represents the cost of merging clusters CR1 and CR2, while m_j refers to the cluster j center, and the number of values in the cluster is denoted by n_j .

The Euclidean distance equation (3.3)

$$\text{Dist} (x1, y1) = \sqrt{\sum (y1_i - x1_i)^2} \quad \dots \quad (3.3)$$

Algorithm 3.3 shows the Scheduling Algorithm based **agglomerative clustering** that will be executed during the first period of each round.

Algorithm 3.3: Scheduling based agglomerative clustering

Input: $V = v_1 \dots, v_{NS}$; **NS:** number of sensors

Output: **S:** the schedule of sensor nodes

- 1:** for $j \leftarrow 1$ to NS **do**
- 2:** $cen_j \leftarrow \{v_j\}$;
- 3:** **end**
- 4:** $Cl \leftarrow \{cen_1, \dots, cen_p\}$;
- 5:** **while** Length (Cl) > 1 **do**
- 6:** $dist(cen_i, cen_j) \leftarrow warddist(cen_i, cen_j), \forall cen_i, cen_j \in Cl$;
- 7:** Add merged clusters (cen_i, cen_j) to Cl ;
- 8:** $Cl \leftarrow Cl \setminus \{cen_i, cen_j\}$ // Eliminate cen_i and cen_j from Cl ;
- 9:** **end**
- 10:** $\{Clus^1, Clus^2, \dots, Clus^k\} \leftarrow$ Bring Clusters (Cl, K);
- 11:** **For** each cluster $j \in Clus^j$ **do** // $j \in K$
- 12:** Select S_i has higher remaining energy to be Active for each cluster $j, S_i \leftarrow 1$;
- 13:** Set S_i in each cluster to be Sleep if $i \neq 1$, where $S_i \leftarrow 0$;
- 14:** **end for**
- 15:** **return S;**

The steps 1-10 are responsible for clustering the sensor nodes based on their variances into K groups $Clus1, Clus2, \dots, Clusk$. The steps 11-14 select one sensor node from each cluster to represent the cluster in the next period as active node. This node is selected based on the higher remaining energy node in the cluster. The gateway sends a notification packet based on the the schedule of sensor nodes S to notify the sensor nodes to stay active or go to sleep mode. In the second period of each round, only the active nodes will gather and sent the compressed data to the gateway node.

B. Scheduling based k-means Algorithm

In this experiment, we used the k-means clustering algorithm. Since it is one of the easiest unsupervised learning methods used for data clustering. It is a simple and easy approach to group the data with a predefined number of clusters (k). is employed by the proposed scheduling algorithm by which the sensor nodes are clustered according to their correlated data. High correlation of the spatial data between readings of the sensor nodes in the WSN can be exploited particularly in dense sensor networks, where nearby sensor nodes are supposed to sense similar data. Therefore, the most similar sensor data are clustered in the same cluster.

The clustering algorithm is applied to conduct the scheduling of sensor nodes. After grouping the sensors into clusters, each cluster will contain the sensors whose readings are the same. The role of scheduling here is to keep only one sensor transmitting the data, and the others of the sensors go into sleep mode. The scheduling process here is by simulation process, sending a notification (0 or 1), representing node states (i.e., **active/sleep**) to each sensor node for the phase of the sensing in the current period. so, 0 refers to sleep mode in the next period, 1 refers to active mode in the next period. [Algorithm 3.4](#) shows the Scheduling Algorithm-based k-means Algorithm that will be executed during step 2 in each first period of each round.

Algorithm 3.4 Scheduling based k-means Algorithm

Require: D : vector of data values, k : number of clusters

Ensure: S : the schedule of sensor nodes.

1: $tm \leftarrow 0$

2: Initialize k centers: $m_1^{tm}, \dots, m_k^{tm}$

3: Repeat

4: $tm \leftarrow tm + 1$

```

5:  $Cl_j \leftarrow \emptyset // \forall j \in k$ 
6: for each  $x_j \in D$  do
7:    $s = \arg\text{Min}_i \{ \|x_j - m_i^{tm}\| \}$  // allocate  $x_j$  to the nearest point center
8:    $Cl_s \leftarrow Cl_s \cup \{x_j\}$ 
9: end for
10: for each  $i \in k$  do
11:    $m_i^{tm} \leftarrow \frac{1}{|Cl_i|} \sum_{x_j \in Cl_i} x_j$ 
12: end for
13: Until convergence
14:  $\{Cl^1, Cl^2, \dots, Cl^k\} \leftarrow \text{Bring Clusters } (Cl, K);$ 
15: For each cluster  $j \in Cl^j$  do //  $j \in K$ 
16:   Select  $S_i$  has higher remaining energy to be Active for each cluster  $j$ ,  $S_i \leftarrow 1$ ;
17:   Set  $S_i$  in each cluster to be Sleep if  $i \neq 1$ , where  $S_i \leftarrow 0$ ;
18: end for
19: return S

```

Algorithm 3.4 is executed at the gateway level to provide the best sensor node schedule for the next period during the present round. After that, the sink (or Gateway) sends notification packets to all sensor nodes in the WSN to inform them who will stay active and who will go to sleep in the next period. Active sensors continue monitoring the area of interest to send their data to the sink next period. The gateway uses node states vector S (0 for sleep, 1 for active) which has been found in Algorithm 3.4 to decide the nodes that should keep sending data to sink and the nodes that should switch into sleep mode until the beginning of the next round.

Algorithm 3.5 explains steps of the proposed approach at both levels (sensors and gateway) levels.

Algorithm 3.5 Proposed Scheduling technique

```

1: For each round  $r$  do //  $r \leftarrow 1, 2, \dots, R$ 
2:   Set each  $S_i$  to 1 (1 Active, 0 Sleep) // Wakeup all sensors in the list  $S$ 
3:   For each period  $p$  do //  $p \leftarrow 1, \dots, P // P \leftarrow 2$ 
4:     For each  $i$  in  $S$  do //  $i \in NS$ ,  $NS$ = number of sensor nodes
5:       If  $S_i$  is Active then
6:          $Data_i^p \leftarrow \text{Collect}()$  // collect data of  $S_i$  for period  $p$ 
7:          $ComFile_i^p \leftarrow \text{CompressFunc}(Data_i^p)$ 
8:         Send  $ComFile_i^p$  to the Gateway
9:         Update remaining energy of  $S_i$ 
10:       Endif
11:     End for
12:   If  $p = 1$  then
13:     For each  $i$  in  $S$  do
14:        $Data_i^p \leftarrow \text{DecompressFunc}(ComFile_i^p)$ 
15:        $V_i^p \leftarrow \text{Variance}(Data_i^p)$ 
16:     End for
17:      $S \leftarrow \text{Scheduling Algorithm based Clustering}(V, NC)$ 
18:     Send Notification Packet to every  $S_i$ ,  $i \in NS$  to inform her to stay Active
    or Sleep
19:     Update remaining energy of  $S_i$ 
20:   Endif
21:   Send collected data during period  $p$  to the Sink.
22: End for
23: End for

```

The steps (5-11) are applied in sensors levels, while steps (12-20) are applied in Gateway levels.

- ❖ The CompressFunc uses either [Algorithm 3.1](#) or [Algorithm 2.1](#).
- ❖ The DecompressFunc uses either [Algorithm 3.2](#) or [Algorithm 2.2](#).

3.3 Summary

In this chapter, the proposed approach is explained, which is scheduling sensor nodes that search the correlation among the nodes' data sent to the sink to make sleep/active scheduling which eventually conserves the network energy . Our proposed technique operates into two levels: sensor nodes and the gateway. The nodes catch the sensed data and reduce them using either Modified Run Length Encoding (MRLE) or Huffman Encoding (HE). In addition, the scheduling technique implemented is based on either the agglomeration clustering or k-means methods to provide the optimal schedule of the sensor nodes that remain active in the next period at the gateway layer.

Chapter Four
Simulation Results and
Discussions

Chapter Four

Simulation Results and Discussions

4.1 Introduction

This chapter explains the results of simulating and testing the proposed approach explained in Chapter 3 and evaluating its performance by performing several experiments based on different performance metrics. In addition, to verify the efficiency of the suggested approach we perform a comparison with another scheduling method under the same performance metrics. The results of the simulation show that the suggested approach is very efficient in saving energy and extending the lifetime of the network over other scheduling methods.

4.2 Simulation Framework

This section shows the simulation findings using a custom simulator-based python programming language. A real sensed data from the sensor network has been utilized in Intel Berkeley Research Lab [77]. This simulation considered the positions of sensor nodes and the related data set files of each sensor node. Each sensor will send its data periodically and wait for the message from the gateway to inform it to stay active or sleep for the next period. The simulation continues until all the data in the data file of sensor nodes are finished.

The main goal of the conducted experimental results is to show the effectiveness of the suggested approach. [Figure 4.1](#) displays the deployment of the sensors in the Intel Laboratory. The sensed data from some nodes are perhaps lost or truncated (See the yellow mark in [Figure 4.1](#)). We are considered only the temperature in this simulation for the sake of simplicity.

It is supposed that all the sensor devices sent their data to the gateway located at the laboratory center.

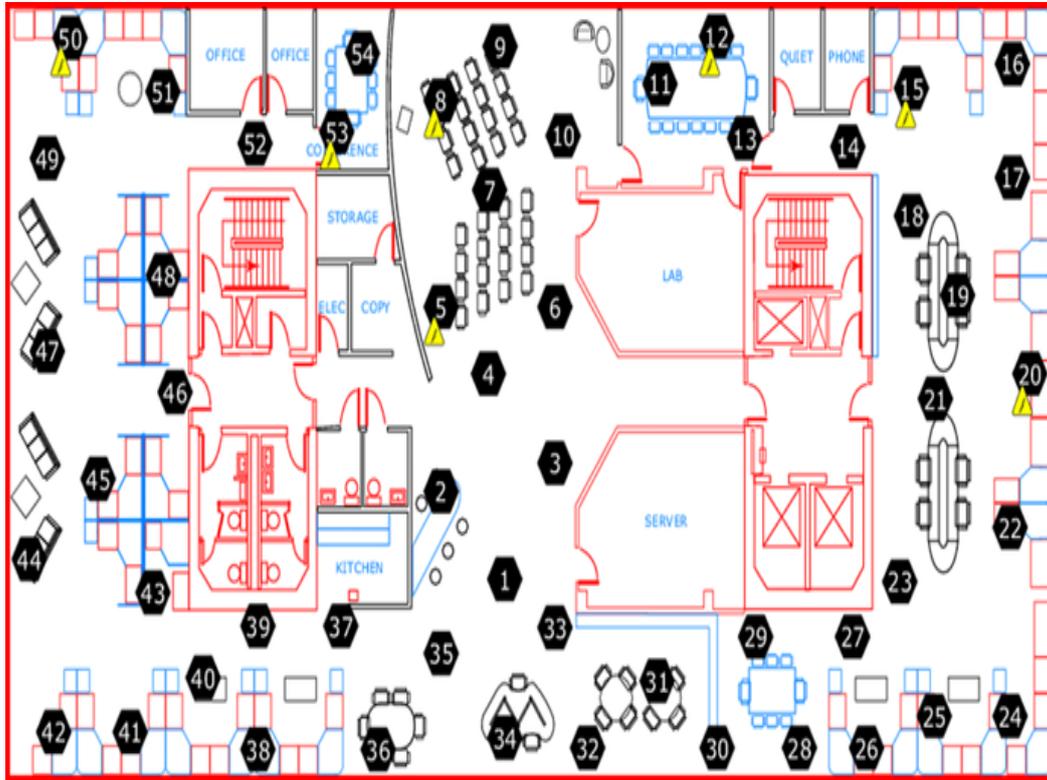


Figure 4.1: The Deployment of the Sensors in the Intel Laboratory

In the simulation, the proposed approach employed the energy consumption model introduced in section 2.9. This model took into account only the consumed energy of transmitting and receiving.

The proposed technique compared with another method called the PFF (Prefix-Frequency Filtering) technique using the same dataset and the same simulation parameters to ensure the effectiveness of the proposed approach. The suggested techniques performance has been assessed using a variety of performance indicators. such as Active Sensor Ratio, Energy Consumption in Joules, Transmitted Data Ratio in KB, and Percentage of Lost Data. The values of the parameters utilized in the simulations are shown in Table 4.1.

Table 4.1: Parameters Setting

Parameter	Value
Number of Sensors (NS)	47
Size of Period (P)	20, 50, and 100
Number of Clusters (NC)	30
Threshold α	0.03, 0.05, and 0.07
E_{elec}	$50 * 10^{-9}$
β_{amp}	$100 * 10^{-12}$

4.3 Performance Evaluation and Analysis

In this section, the efficiency of the proposed approach is evaluated using various measures. The key aim of these experiments is to show that our approach will effectively produce interesting results for energy conservation at sensor nodes in WSN.

4.3.1 Selecting the Number of Clusters (NC)

In this experiment, we explain the effect of the Number of Clusters (NC) on the performance of the proposed approach. Table 4.2 shows the effect (NC) on the performance of the scheduling approach. This experiment is performed on the network of the sensor of 47 sensors (In Fact, the total number of sensors is 54, but 7 sensors of them were perhaps lost or truncated, so we based our experiments on 47 sensors only), $p = 50$, and $\alpha = 0.02$. where refers (Active Sensor Ratio (ASR), Transmitted Data Ratio (TD), Energy Consumption (EC), Percentage of Lost Data (PLD)).

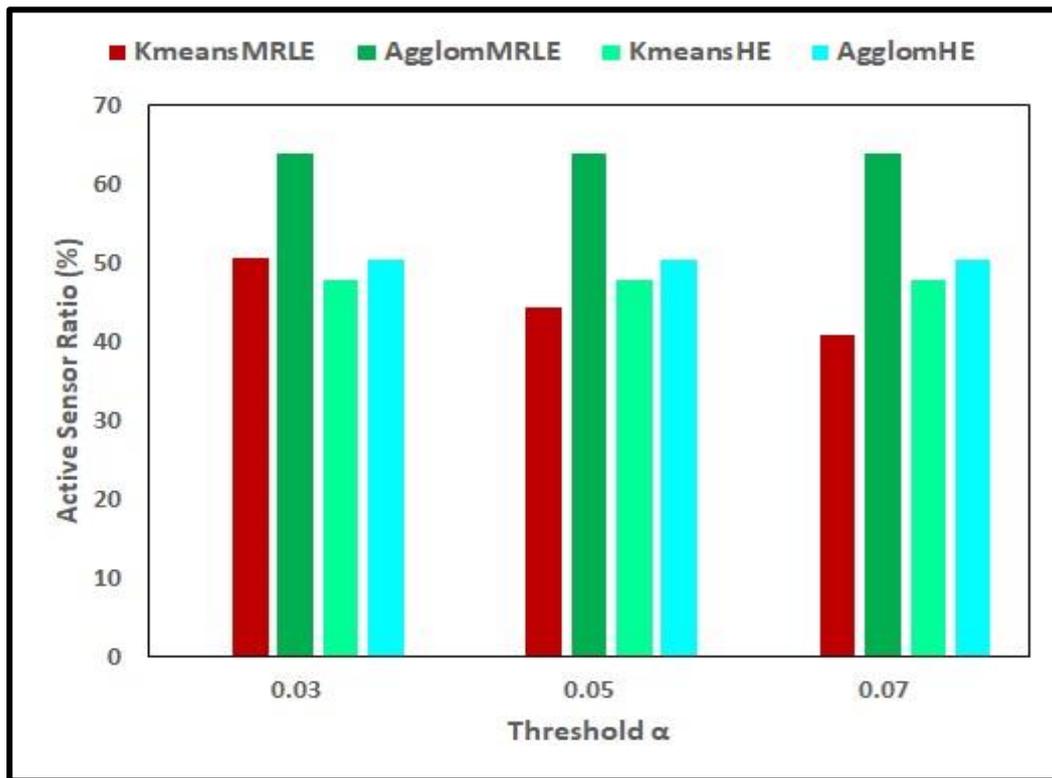
Table 4.2: Number of Clusters (NC) vs ASR, TD, EC, and PLD

NC	ASR	TD	EC	PLD
10	20.86	4.00	0.0024	16.02
15	31.14	4.40	0.0027	10.07
20	41.41	4.79	0.0029	6.17
25	51.38	5.15	0.0031	3.54
30	60.54	5.46	0.0033	2.16
35	68.73	5.74	0.0035	1.27
40	75.08	5.94	0.0036	0.76

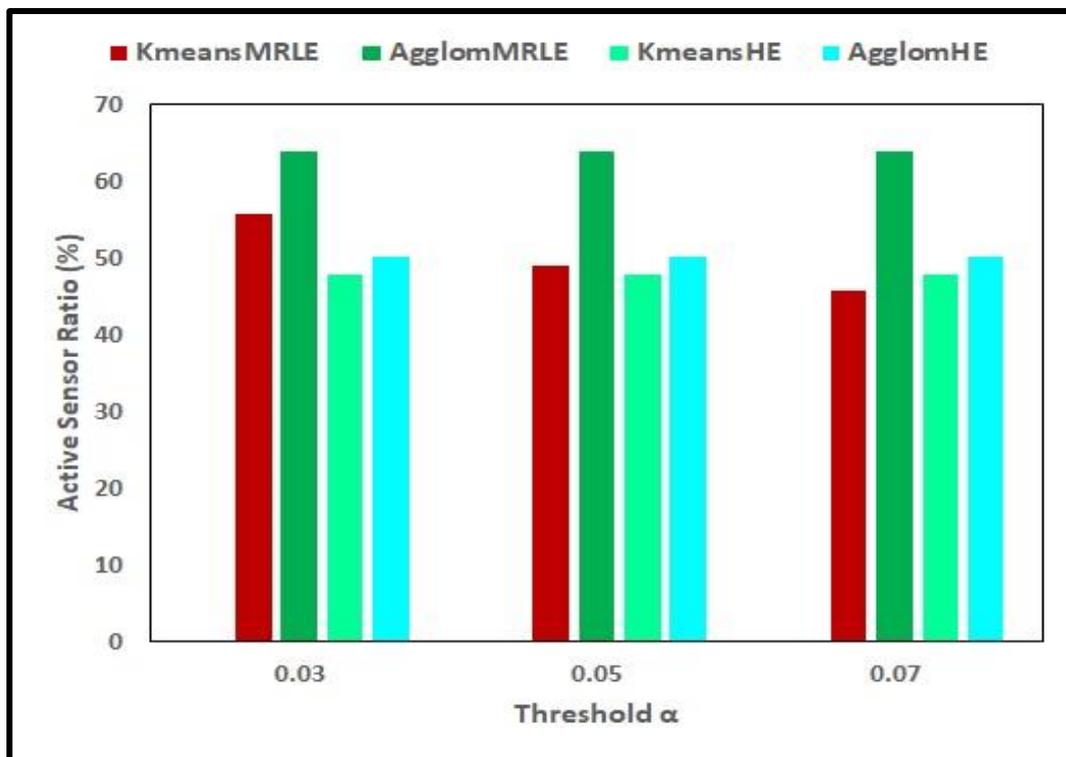
As shown in Table 4.2, when the NC increased the transmitted data to the sink increased since the number of active sensors increased. Hence, the consumed energy increased, and the total number of lost measures decreased. Therefore, it is important to select several clusters that balance the consumed energy and the lost data. Experimentally we select NC equal to the 30 and fixed for all experiments.

4.3.2 Active Sensor Ratio (ASR)

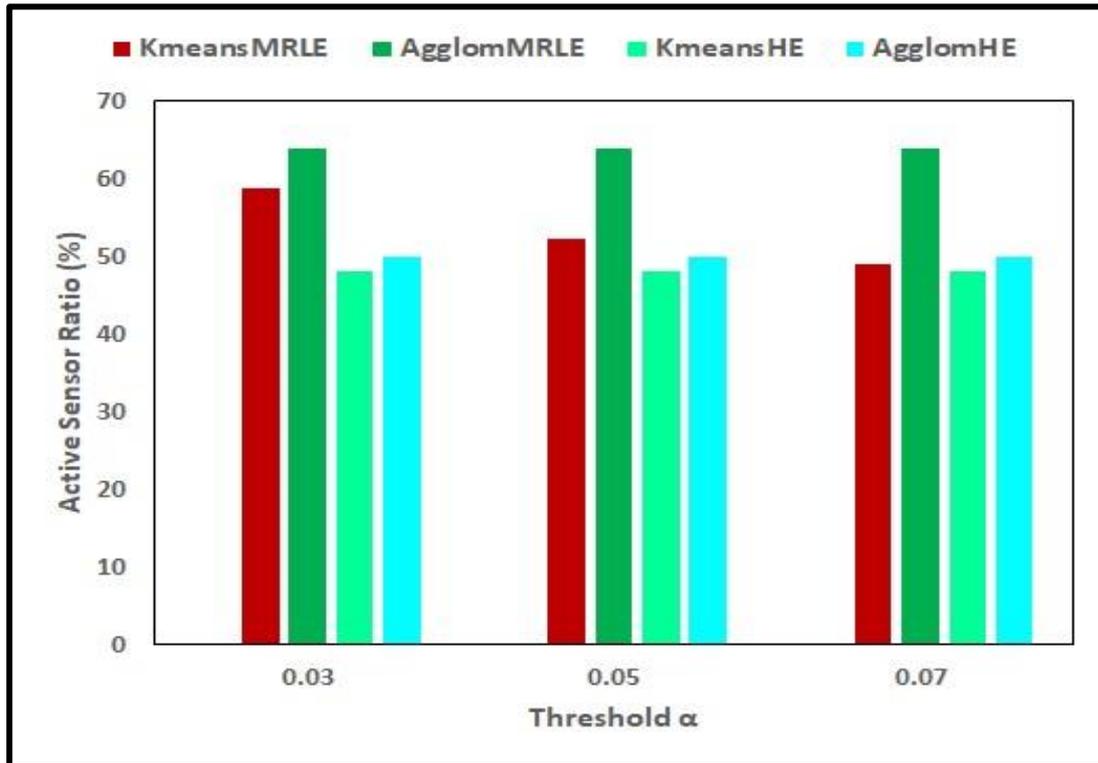
One key performance metric is to reduce the number of active sensors that take on the sensing task during the next period while maintaining an acceptable data accuracy at the gateway node. The active sensor ratio is shown in Figure 4.2 for various data sizes and α values.



[a] p=20



[b] p=50



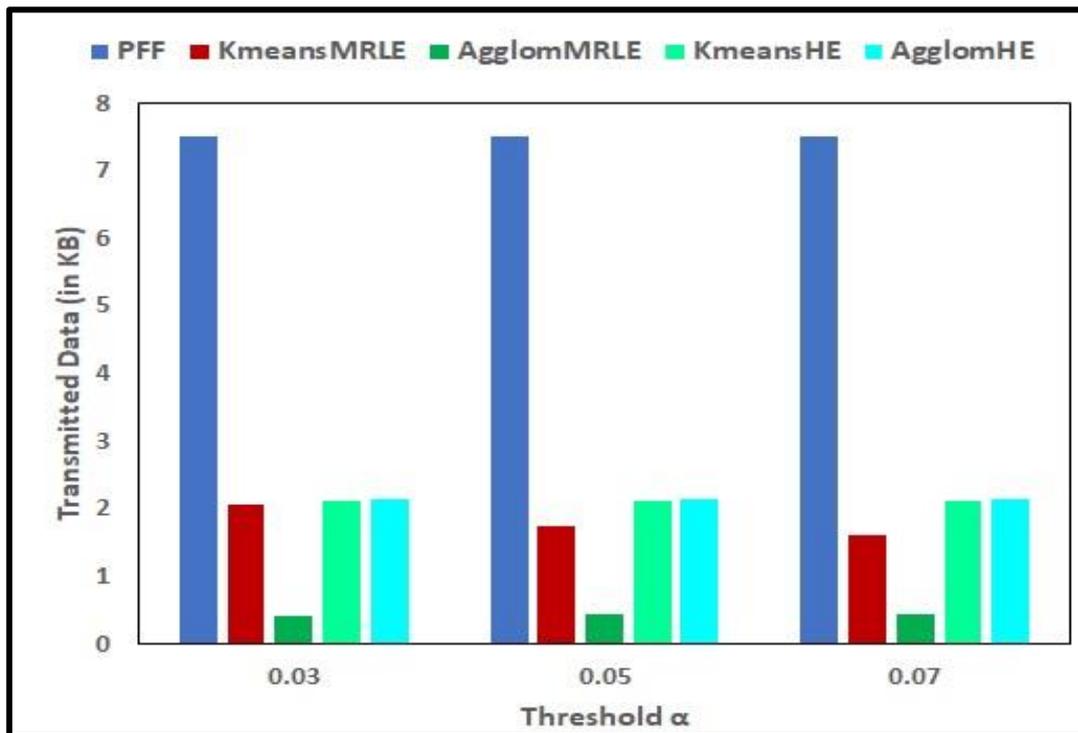
[c] $p=100$

Figure 4.2: Active Sensor Ratio

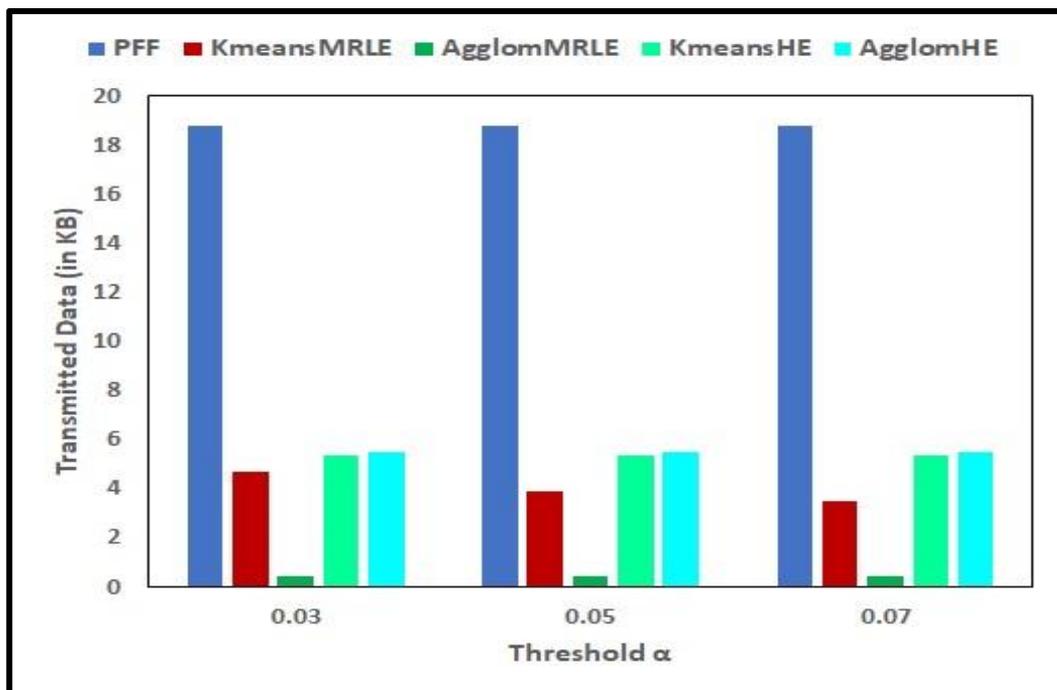
As shown in the results of [Figure 4.2](#), The scheduling approach based on (KmeansHE) activates a lower number of sensor nodes in most data sizes per period. The results of the scheduling approach (AgglomHE) are competitive and provide a lower number of active nodes compared with rest approaches except (KmeansHE) . The scheduling approach with lossy compression can increase the activated sensor nodes due to the high reduction in the data arrived at the gateway.

4.3.3 Transmitted Data Ratio (TDR)

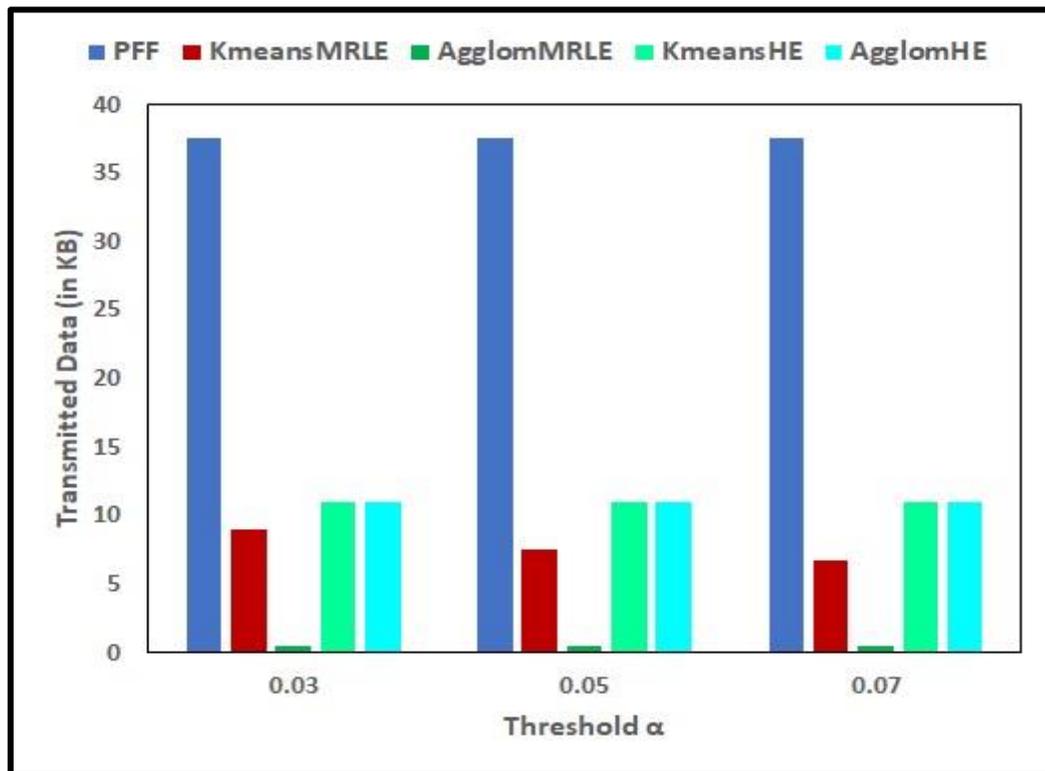
In [Figure 4.3](#), we show the effect of the parameters P and α on the amount of transmitted data to the gateway node per period. [Figure 4.3](#) depicts the ratio of transmitted data to gateway per period.



[a] P=20



[b] P=50



[c] P=100

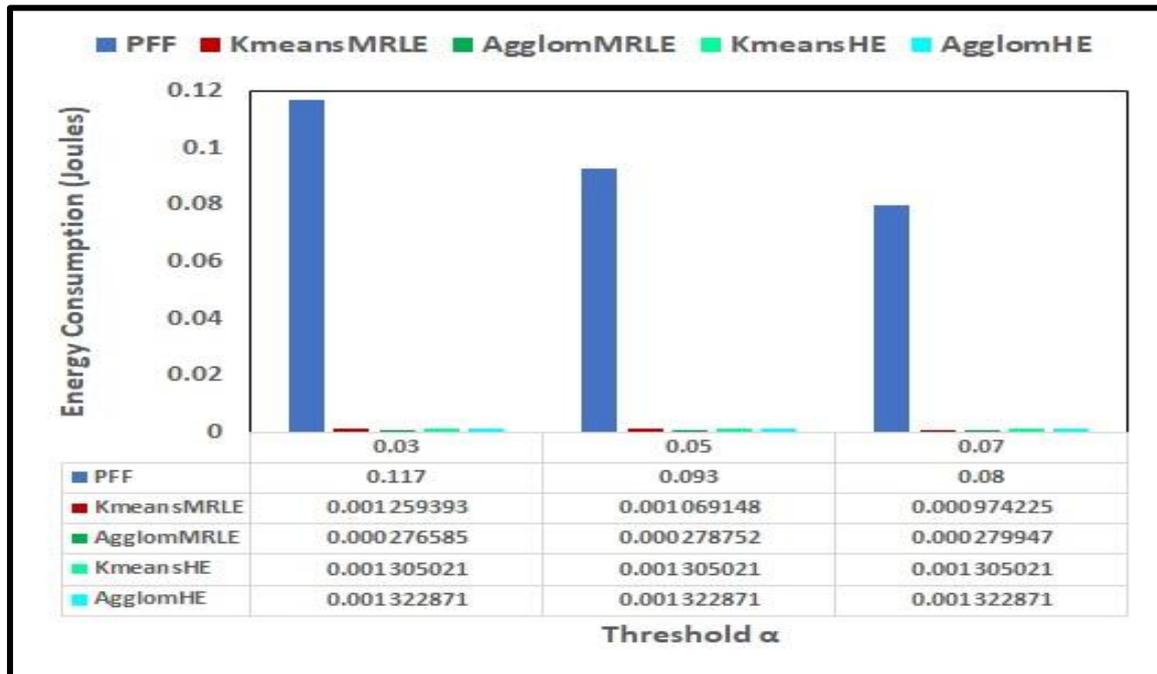
Figure 4.3: Transmitted Data Ratio

The scheduling approaches i.e. KmeansMRLE and AgglomMRLE reduced the ratio of sent data higher than another approach. The AgglomMRLE decreased the transmitted data from 94.3 up to 98.8 compared to the PFF. The KmeansMRLE decreased the transmitted data from 75.9% up to 78.7% compared to the PFF.

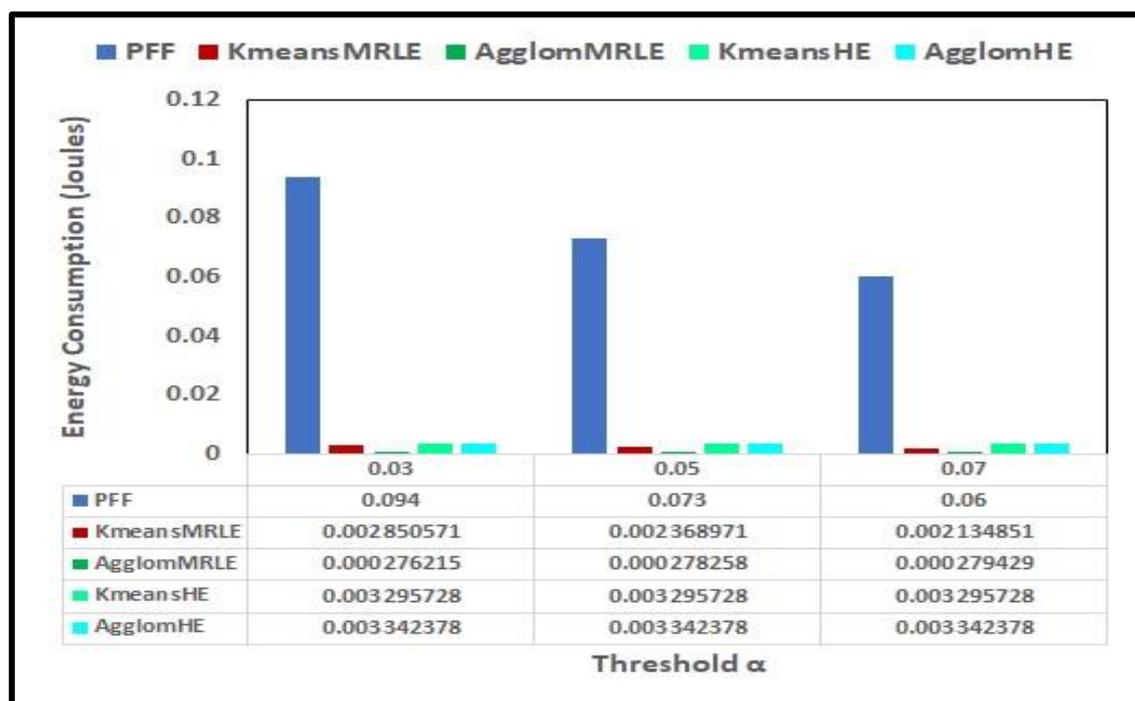
This is due to deactivating the sensor nodes that transmit the same data to the gateway node during each round. Furthermore, integrate the scheduling approach with a lossy compression method to further compress the sent data by each sensor node to the gateway.

4.5.4 Energy Consumption (EC)

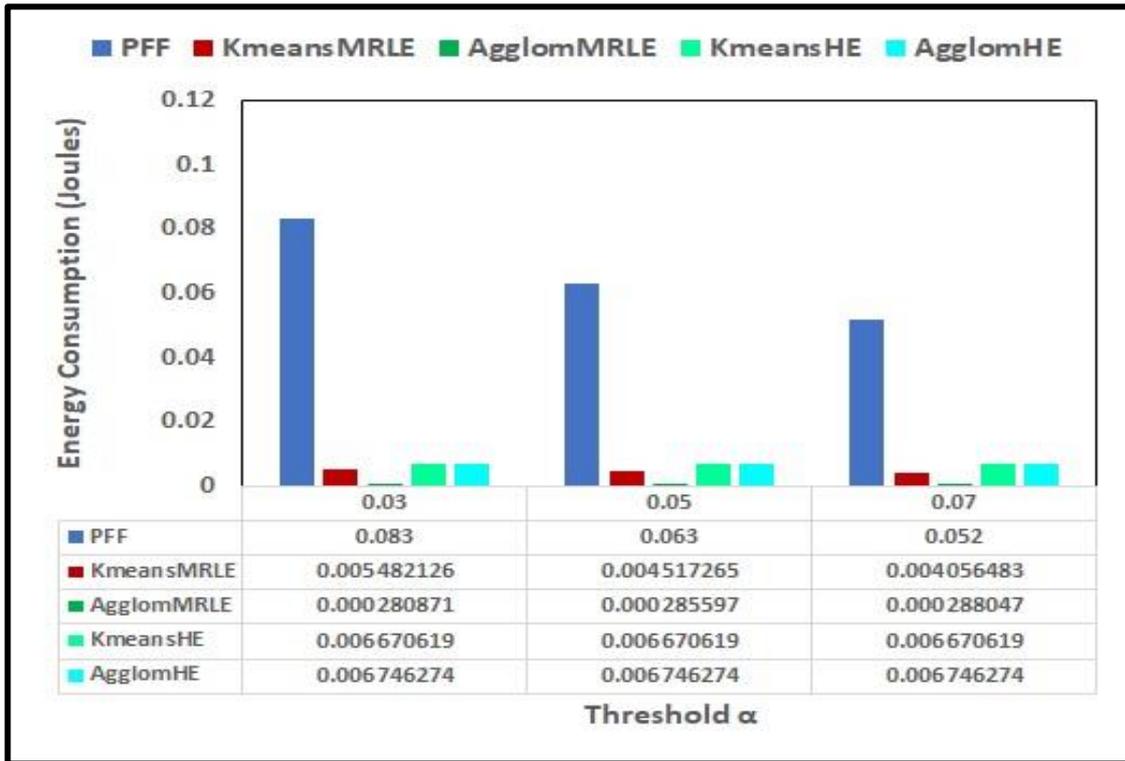
To show the effectiveness of the suggested approach, EC measure by the sensor node is computed. Figure 4.4 demonstrates how the proposed approach outperforms the PFF method in terms of energy usage.



[a] P=20



[b] P=50



[c] P=100

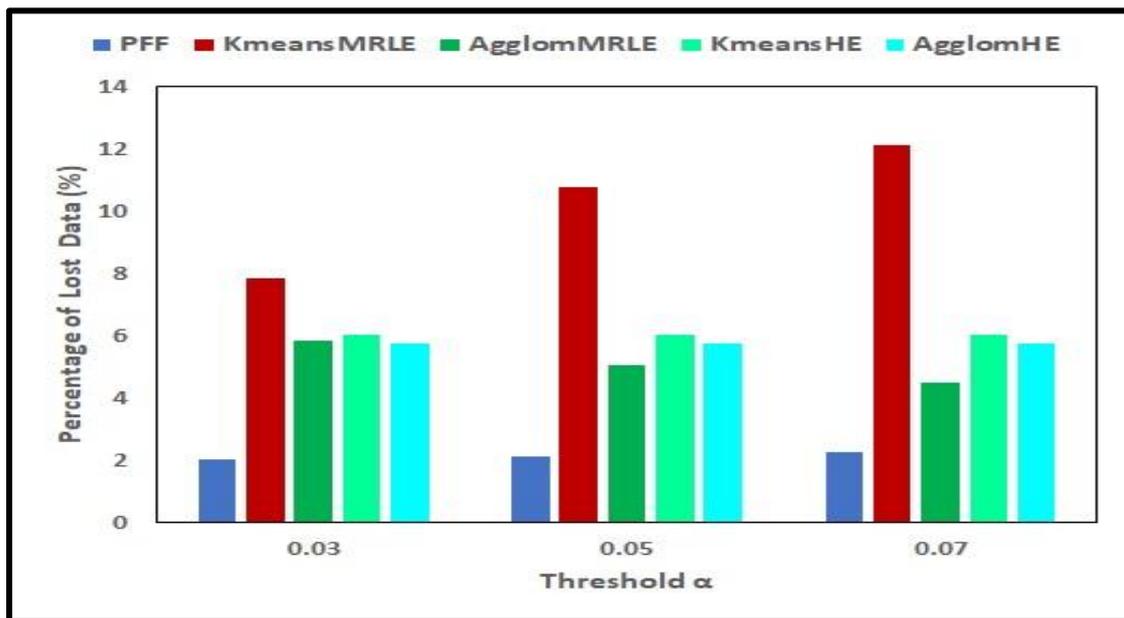
Figure 4.4. Energy Consumption

The scheduling approaches (KmeansMRLE and AgglomMRLE) consumed less energy compared with other approaches due to transmitting fewer data to the gateway. The AgglomMRLE decreased the consumed energy from 99.5% up to 99.8% compared to the PFF. The KmeansMRLE decreased the consumed energy from 95.3% up to 98.1% compared to the PFF.

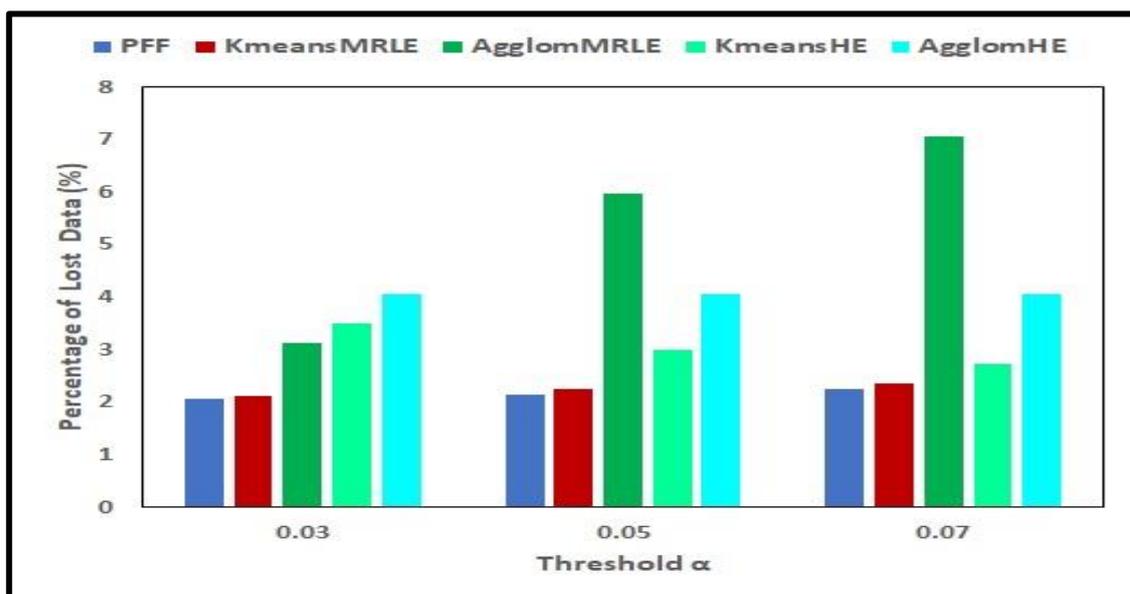
This enhanced result in saving the energy of the WSN comes from the effectiveness of the proposed scheduling approach that activates as few sensor nodes as possible per period while maintaining a good data quality level at the gateway node. In addition, the compression approaches that are utilized further reduced the sent data to the gateway and led to saving the energy of sensor nodes in the IoT network.

4.5.5 Percentage of Lost Data (PLD)

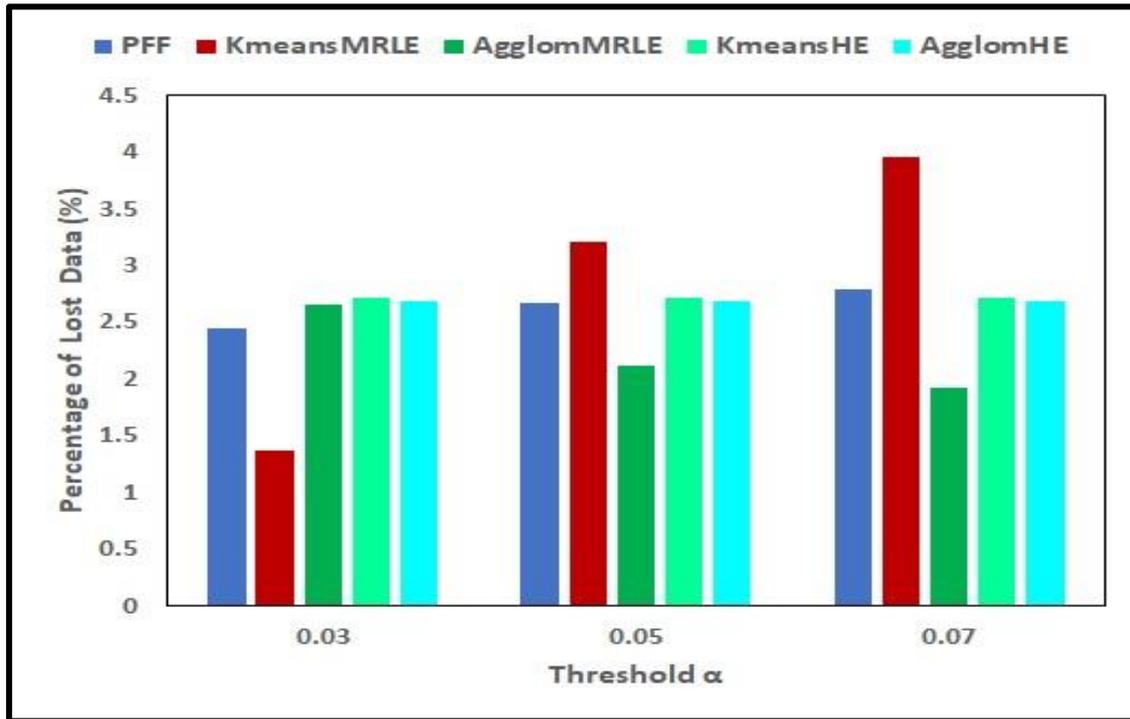
In this experiment, the data loss ratio of the received data at the gateway is studied. It is important to design an energy-saving technique balance between the consumed energy and the percentage of data loss so that no significant data loss affects the quality of data at the gateway. Figure 4.5 shows the Percentage of Lost Data.



[a] P= 20



[b] P= 50



[c] P= 100

Figure 4.5. Percentage of Lost Data

Figure 4.5 shows that the proposed approach relatively lost higher data with P=20 and 50 as compared to the PFF approach. Because our approach highly reduces the sent data to the gateway while saving energy and the suitable quality of data at the gateway. Our method produced a higher data accuracy with P=100 in comparison with PFF despite the high reduction of transmitted data to the gateway .

The AgglomMRLE decreased the percentage of lost data from 5.3% up to 21.1 % compared to the PFF. Furthermore, scheduling approaches with HE gives a better performance in decreasing the lost data ratio compared with PFF when P=100. Due to impact, the lossless compression method assisted in increasing the accuracy of the data at the sink.

Therefore, the suggested approach preserves the original data and considerably decreases the rate of data loss significantly while optimizing the lifetime of the IoT network.

4.5 Summary

This chapter explains the evaluation results of the proposed technique using a custom simulator-based python programming language as well as real data of a sensor network. Within this chapter, the suggested technique's performance has been assessed using a variety of performance indicators, such as Active Sensor Ratio (**ASR**), Energy Consumption (**EC**) in Joules, Transmitted Data Ratio (**TDR**) in KB, and Percentage of Lost Data (**PLD**). The comparison results show that our proposed technique outperforms the PFF technique in terms of several performance metrics like the active sensor ratio, data sent ratio, data loss ratio, energy-saving, and energy consumptions.

Chapter Five
Conclusions and Suggestions for
Future Works

Chapter Five

Conclusions and Suggestions for Future Works

5.1 Conclusions

Based on the conducted results of the proposed approach, some points can be concluded and can be summarized as follows:

1. This search thesis suggests an Energy-efficient scheduling approach using clustering and encoding in WSNs.
2. The proposed scheduling techniques provides decreased the number of active sensor nodes during monitoring the area of interest using an energy-efficient scheduling way saving the batteries power of the sensor nodes in WSNs.
3. The proposed approach uses encoding approaches like HE and MRLE to further reduce the transmitted data to the sink. The MRLE method provides complete results when it is employed at the sensor node compared with other methods.
4. When the NC increased the transmitted data to the sink increased, the consumed energy increased, and the total number of lost measures decreased. Therefore, it is important to select several clusters that balance the consumed energy and the lost data.
5. The AgglomMRLE decreased the transmitted data from 94.3% up to 98.8% compared to the PFF. The KmeansMRLE decreased the transmitted data from 75.9% up to 78.7% compared to the PFF.
6. The AgglomMRLE decreased the consumed energy from 99.5% up to 99.8% compared to the PFF. The KmeansMRLE decreased the consumed energy from 95.3% up to 98.1% compared to the PFF.
7. The proposed approach loses a little bit higher data with $P=20$ and 50 as compared to the PFF approach. Our method produced a higher

data accuracy with $P=100$ in comparison with PFF despite the high reduction of transmitted data to the gateway . The AgglomMRLE decreased the percentage of lost data from 5.3% up to 21.1 % compared to the PFF. Furthermore, scheduling approaches with HE introduced a better performance by decreasing the lost data ratio compared with PFF in the size $P=100$ due to utilizing the lossless compression method that assisted in increasing the accuracy of the data at the sink.

5.2 Future Works

1. We plan to provide the schedules for more than one period.
2. We plan to implement the presented work on a cluster-based network by executing the scheduling algorithm inside each cluster head.
3. The deep learning approaches can be exploited to schedule the sensor nodes either using grouping or optimization.
4. The encoding methods can be improved to further decrease the transmitted data to the sink.
5. It is planned to employ a data prediction approach at the sink to predict the missing data and increase the accuracy of received data at the sink.
6. Real experiments could be made to assess the efficiency of the proposed technique.

References ...

- [1] F. Yu, C.-C. Chang, J. Shu, I. Ahmad, J. Zhang, and J. M. de Fuentes, "Recent Advances in Security and Privacy for Wireless Sensor Networks 2016," ed: Hindawi, 2017.
- [2] Y. Chen, Z. Wang, T. Ren, Y. Liu, and H. Lv, "Maximizing lifetime of wireless sensor networks with mobile sink nodes," *Mathematical Problems in Engineering*, vol. 2014, 2014.
- [3] R. E. Mohamed, A. I. Saleh, M. Abdelrazzak, and A. S. Samra, "Survey on wireless sensor network applications and energy efficient routing protocols," *Wireless Personal Communications*, vol. 101, no. 2, pp. 1019-1055, 2018.
- [4] M. Abujubbeh, F. Al-Turjman, and M. Fahrioglu, "Software-defined wireless sensor networks in smart grids: An overview," *Sustainable Cities and Society*, vol. 51, p. 101754, 2019.
- [5] Z. Al Aghbari, A. M. Khedr, W. Osamy, I. Arif, and D. P. Agrawal, "Routing in wireless sensor networks using optimization techniques: A survey," *Wireless Personal Communications*, vol. 111, no. 4, pp. 2407-2434, 2020.
- [6] A. Shahraki, A. Taherkordi, Ø. Haugen, and F. Eliassen, "Clustering objectives in wireless sensor networks: A survey and research direction analysis," *Computer Networks*, vol. 180, p. 107376, 2020.
- [7] S. A. Dehkordi, K. Farajzadeh, J. Rezazadeh, R. Farahbakhsh, K. Sandrasegaran, and M. A. Dehkordi, "A survey on data aggregation techniques in IoT sensor networks," *Wireless Networks*, vol. 26, no. 2, pp. 1243-1263, 2020.
- [8] Y. Tripathi, A. Prakash, and R. Tripathi, "A sleep scheduling based cooperative data transmission for wireless sensor network," *International Journal of Electronics*, pp. 1-22, 2021.

- [9] S. Pushpalatha and K. Shivaprakasha, "Energy-efficient communication using data aggregation and data compression techniques in wireless sensor networks: A survey," in *Advances in communication, signal processing, VLSI, and embedded systems*: Springer, 2020, pp. 161-179.
- [10] K. Sohraby, D. Minoli, and T. Znati, *Wireless sensor networks: technology, protocols, and applications*. John wiley & sons, 2007.
- [11] J. M. Bahi, A. Makhoul, M. J. A. Medlej, and S. W. Networks, "A two tiers data aggregation scheme for periodic sensor networks," vol. 21, no. 1, 2014.
- [12] J. Feng and H. Zhao, "Energy-balanced multisensory scheduling for target tracking in wireless sensor networks," *Sensors*, vol. 18, no. 10, p. 3585, 2018.
- [13] K. P. Mhatre and U. P. Khot, "Energy Efficient Opportunistic Routing with Sleep Scheduling in Wireless Sensor Networks," *Wireless Personal Communications*, vol. 112, no. 2, pp. 1243-1263, 2020.
- [14] P. S. Banerjee, S. N. Mandal, D. De, and B. Maiti, "RL-sleep: Temperature adaptive sleep scheduling using reinforcement learning for sustainable connectivity in wireless sensor networks," *Sustainable Computing: Informatics and Systems*, vol. 26, p. 100380, 2020.
- [15] M. N. Khan *et al.*, "Improving energy efficiency with content-based adaptive and dynamic scheduling in wireless sensor networks," *IEEE Access*, vol. 8, pp. 176495-176520, 2020.
- [16] W. Feng *et al.*, "Joint energy-saving scheduling and secure routing for critical event reporting in wireless sensor networks," *IEEE Access*, vol. 8, pp. 53281-53292, 2020.

- [17] M. N. Khan, H. U. Rahman, and M. Z. Khan, "An energy efficient adaptive scheduling scheme (EASS) for mesh grid wireless sensor networks," *Journal of Parallel and Distributed Computing*, vol. 146, pp. 139-157, 2020.
- [18] G. Yang, X.-W. Wu, Y. Li, and Q. Ye, "Energy efficient protocol for routing and scheduling in wireless body area networks," *Wireless Networks*, vol. 26, no. 2, pp. 1265-1273, 2020.
- [19] N. Mahendran, S. Shankar, and T. Mekala, "EMA-PRBDS: efficient multi-attribute packet rank based data scheduling in wireless sensor networks for real-time monitoring systems," *International Journal of Electronics*, vol. 107, no. 6, pp. 879-898, 2020.
- [20] R. Sinde, F. Begum, K. Njau, and S. J. S. Kaijage, "Refining network lifetime of wireless sensor network using energy-efficient clustering and DRL-based sleep scheduling," vol. 20, no. 5, p. 1540, 2020.
- [21] C. S. Raghavendra, K. M. Sivalingam, and T. Znati, *Wireless sensor networks*. Springer, 2006.
- [22] A. B. Noel, A. Abdaoui, T. Elfouly, M. H. Ahmed, A. Badawy, and M. S. Shehata, "Structural health monitoring using wireless sensor networks: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 3, pp. 1403-1423, 2017.
- [23] M. P. Đurišić, Z. Tafa, G. Dimić, and V. Milutinović, "A survey of military applications of wireless sensor networks," in *2012 Mediterranean conference on embedded computing (MECO)*, 2012, pp. 196-199: IEEE.
- [24] B. Rashid and M. H. Rehmani, "Applications of wireless sensor networks for urban areas: A survey," *Journal of network and computer applications*, vol. 60, pp. 192-219, 2016.

- [25] S. Chitnis, N. Deshpande, and A. Shaligram, "An investigative study for smart home security: Issues, challenges and countermeasures," *Wireless Sensor Network*, vol. 8, no. 04, p. 61, 2016.
- [26] M. Y. Aalsalem, W. Z. Khan, W. Gharibi, M. K. Khan, and Q. Arshad, "Wireless Sensor Networks in oil and gas industry: Recent advances, taxonomy, requirements, and open challenges," *Journal of network and computer applications*, vol. 113, pp. 87-97, 2018.
- [27] S. Mahfoudh, P. Minet, and I. Amdouni, "Energy efficient routing and node activity scheduling in the OCARI wireless sensor network," *Future Internet*, vol. 2, no. 3, pp. 308-340, 2010.
- [28] B. A. Akgül, M. F. Hasoğlu, and B. J. W. S. N. Haznedar, "Investigation and Implementation Ultra-Low Power PIC-Based Sensor Node Network with Renewable Energy Source and Decision-Making Unit," vol. 10, no. 2, pp. 41-58, 2018.
- [29] I. Khemapech, I. Duncan, and A. Miller, "A survey of wireless sensor networks technology," in *6th Annual Postgraduate Symposium on the Convergence of Telecommunications, Networking and Broadcasting*, 2005, vol. 13.
- [30] F. Engmann, F. A. Katsriku, J.-D. Abdulai, K. S. Adu-Manu, and F. K. Banaseka, "Prolonging the lifetime of wireless sensor networks: a review of current techniques," *Wireless Communications and Mobile Computing*, vol. 2018, 2018.
- [31] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. J. C. n. Cayirci, "Wireless sensor networks: a survey," vol. 38, no. 4, pp. 393-422, 2002.
- [32] M. Pike, N. M. Mustafa, D. Towey, and V. Brusic, "Sensor networks and data management in healthcare: Emerging technologies and new challenges," in *2019 IEEE 43rd Annual Computer Software and*

- Applications Conference (COMPSAC)*, 2019, vol. 1, pp. 834-839: IEEE.
- [33] T. Arampatzis, J. Lygeros, and S. Manesis, "A survey of applications of wireless sensors and wireless sensor networks," in *Proceedings of the 2005 IEEE International Symposium on, Mediterrean Conference on Control and Automation Intelligent Control, 2005.*, 2005, pp. 719-724: IEEE.
- [34] S. Prasanna and S. Rao, "An overview of wireless sensor networks applications and security," *International Journal of Soft Computing and Engineering (IJSCE)*, vol. 2, no. 2, pp. 2231-2307, 2012.
- [35] S. J. Ramson and D. J. Moni, "Applications of wireless sensor networks—A survey," in *2017 international conference on innovations in electrical, electronics, instrumentation and media technology (ICEEIMT)*, 2017, pp. 325-329: IEEE.
- [36] H. Sharma, A. Haque, and Z. A. Jaffery, "Maximization of wireless sensor network lifetime using solar energy harvesting for smart agriculture monitoring," *Ad Hoc Networks*, vol. 94, p. 101966, 2019.
- [37] J. Prajapati and S. Jain, "Machine learning techniques and challenges in wireless sensor networks," in *2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT)*, 2018, pp. 233-238: IEEE.
- [38] S. Sharma, R. K. Bansal, and S. Bansal, "Issues and challenges in wireless sensor networks," in *2013 International Conference on Machine Intelligence and Research Advancement*, 2013, pp. 58-62: IEEE.
- [39] K. Gupta and V. Sikka, "Design issues and challenges in wireless sensor networks," *international journal of computer applications*, vol. 112, no. 4, pp. 0975-8887, 2015.

- [40] V. C. Gungor, M. C. Vuran, and O. Akan, "On the cross-layer interactions between congestion and contention in wireless sensor and actor networks," *Ad Hoc Networks*, vol. 5, no. 6, pp. 897-909, 2007.
- [41] K. Vikhyath, S. J. I. J. o. E. Brahmanand, and Technology, "Wireless sensor networks security issues and challenges: A survey," vol. 7, no. 2.33, pp. 89-94, 2018.
- [42] P. J. Sallis, *Wireless Sensor Networks: Insights and Innovations*. BoD–Books on Demand, 2017.
- [43] L. Wang and Y. Xiao, "A survey of energy-efficient scheduling mechanisms in sensor networks," *Mobile Networks and Applications*, vol. 11, no. 5, pp. 723-740, 2006.
- [44] A. Kozłowski and J. J. W. P. C. Sosnowski, "Energy efficiency trade-off between duty-cycling and wake-up radio techniques in IoT networks," vol. 107, no. 4, pp. 1951-1971, 2019.
- [45] R. C. Carrano, D. Passos, L. C. Magalhaes, C. V. J. I. C. S. Albuquerque, and Tutorials, "Survey and taxonomy of duty cycling mechanisms in wireless sensor networks," vol. 16, no. 1, pp. 181-194, 2013.
- [46] V. T. Vu, T. V. Quyen, L. H. Truong, A. M. Le, C. V. Nguyen, and M. T. Nguyen, "Energy efficient approaches in wireless sensor networks," *ICSES Transactions on Computer Networks and Communications*, vol. 6, no. 1, pp. 1-10, 2020.
- [47] A. C. Valera, W.-S. Soh, and H.-P. Tan, "Survey on wakeup scheduling for environmentally-powered wireless sensor networks," *Computer Communications*, vol. 52, pp. 21-36, 2014.
- [48] S. H. Lee and L. J. I. J. o. D. S. N. Choi, "ZeroMAC: Toward a zero sleep delay and zero idle listening media access control protocol

- with ultralow power radio frequency wakeup sensor," vol. 13, no. 8, p. 1550147717716397, 2017.
- [49] M. M. Afsar and M.-H. Tayarani-N, "Clustering in sensor networks: A literature survey," *Journal of Network and Computer Applications*, vol. 46, pp. 198-226, 2014.
- [50] S. Wang and E. Ipek, "Reducing data movement energy via online data clustering and encoding," in *2016 49th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO)*, 2016, pp. 1-13: IEEE.
- [51] T. Gui, C. Ma, F. Wang, J. Li, and D. E. Wilkins, "A novel cluster-based routing protocol wireless sensor networks using spider monkey optimization," in *IECON 2016-42nd Annual Conference of the IEEE Industrial Electronics Society*, 2016, pp. 5657-5662: IEEE.
- [52] A. Shahraki, M. Geitle, and Ø. Haugen, "A comparative node evaluation model for highly heterogeneous massive-scale Internet of Things-Mist networks," *Transactions on Emerging Telecommunications Technologies*, vol. 31, no. 12, p. e3924, 2020.
- [53] J. Lloret, M. Garcia, D. Bri, and J. R. Diaz, "A cluster-based architecture to structure the topology of parallel wireless sensor networks," *Sensors*, vol. 9, no. 12, pp. 10513-10544, 2009.
- [54] A. A. Abbasi and M. Younis, "A survey on clustering algorithms for wireless sensor networks," *Computer communications*, vol. 30, no. 14-15, pp. 2826-2841, 2007.
- [55] M. T. Nguyen, K. A. Teague, and N. Rahnavard, "Inter-cluster multi-hop routing in wireless sensor networks employing compressive sensing," in *2014 IEEE Military Communications Conference*, 2014, pp. 1133-1138: IEEE.
- [56] V. T. Vu *et al.*, "Energy efficient approaches in wireless sensor networks," vol. 6, no. 1, pp. 1-10, 2020.

- [57] D. Ramasur and G. P. Hancke, "A wind energy harvester for low power wireless sensor networks," in *2012 IEEE International Instrumentation and Measurement Technology Conference Proceedings*, 2012, pp. 2623-2627: IEEE.
- [58] N. Elvin and A. Erturk, *Advances in energy harvesting methods*. Springer Science & Business Media, 2013.
- [59] R. Soua and P. Minet, "A survey on energy efficient techniques in wireless sensor networks," in *2011 4th Joint IFIP Wireless and Mobile Networking Conference (WMNC 2011)*, 2011, pp. 1-9: IEEE.
- [60] C. McAnlis and A. Haecky, *Understanding compression: Data compression for modern developers*. " O'Reilly Media, Inc.", 2016.
- [61] S. Mittal and S. J. J. o. S. A. Nag, "A survey of encoding techniques for reducing data-movement energy," vol. 97, pp. 373-396, 2019.
- [62] H. D. Kotha, M. Tummanapally, and V. K. Upadhyay, "Review on Lossless Compression Techniques," in *Journal of Physics: Conference Series*, 2019, vol. 1228, no. 1, p. 012007: IOP Publishing.
- [63] R. Arshad, A. Saleem, and D. Khan, "Performance comparison of Huffman coding and double Huffman coding," in *2016 Sixth International Conference on Innovative Computing Technology (INTECH)*, 2016, pp. 361-364: IEEE.
- [64] N. Dhawale, "Implementation of Huffman algorithm and study for optimization," in *2014 International Conference on Advances in Communication and Computing Technologies (ICACACT 2014)*, 2014, pp. 1-6: IEEE.
- [65] M. G. Omran, A. P. Engelbrecht, and A. J. I. D. A. Salman, "An overview of clustering methods," vol. 11, no. 6, pp. 583-605, 2007.
- [66] A. K. J. P. r. l. Jain, "Data clustering: 50 years beyond K-means," vol. 31, no. 8, pp. 651-666, 2010.

- [67] S. J. Nanda, G. J. S. Panda, and E. computation, "A survey on nature inspired metaheuristic algorithms for partitional clustering," vol. 16, pp. 1-18, 2014.
- [68] C. K. Reddy, *Data Clustering: Algorithms and Applications*. Chapman and Hall/CRC, 2018.
- [69] S. Sreedhar Kumar, M. Madheswaran, B. Vinutha, H. Manjunatha Singh, K. J. I. J. o. E. Charan, and Technology, "A brief survey of unsupervised agglomerative hierarchical clustering schemes," vol. 8, no. 1, pp. 29-37, 2019.
- [70] M. C. Vuran, Ö. B. Akan, and I. F. J. C. N. Akyildiz, "Spatio-temporal correlation: theory and applications for wireless sensor networks," vol. 45, no. 3, pp. 245-259, 2004.
- [71] L. A. Villas, A. Boukerche, D. L. Guidoni, H. A. De Oliveira, R. B. De Araujo, and A. A. Loureiro, "An energy-aware spatio-temporal correlation mechanism to perform efficient data collection in wireless sensor networks," *Computer Communications*, vol. 36, no. 9, pp. 1054-1066, 2013.
- [72] M. Abo-Zahhad, O. Amin, M. Farrag, and A. Ali, "Survey on energy consumption models in wireless sensor networks," *Open Trans. Wirel. Sens. Netw*, vol. 1, no. 1, 2014.
- [73] I. Dietrich and F. J. A. T. o. S. N. Dressler, "On the lifetime of wireless sensor networks," vol. 5, no. 1, pp. 1-39, 2009.
- [74] Y. Chen and Q. J. I. C. l. Zhao, "On the lifetime of wireless sensor networks," vol. 9, no. 11, pp. 976-978, 2005.
- [75] H. Jaleel, A. Rahmani, and M. Egerstedt, "Probabilistic lifetime maximization of sensor networks," *IEEE Transactions on Automatic Control*, vol. 58, no. 2, pp. 534-539, 2012.

- [76] W. B. Heinzelman, A. P. Chandrakasan, and H. J. I. T. o. w. c. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," vol. 1, no. 4, pp. 660-670, 2002.
- [77] S. J. U. I. C. Madden, . <http://berkeley>, intel-research, net/labdata, html, "Intel Berkeley research lab data," ed, 2003.

الملخص

تمثل شبكات الاستشعار اللاسلكية (WSNs) العنصر الأساسي للعديد من التطبيقات في إنترنت الأشياء (IoT) والمدن الذكية في الأيام الحالية والمستقبلية. تؤدي طبيعة نشر هذه الأجهزة في العديد من التطبيقات إلى زيادة حجم البيانات الزائدة عن الحاجة المنقولة عبر شبكة إنترنت الأشياء. يمكن أن يؤدي ذلك إلى تدهور أداء الشبكة واستنفاد طاقة البطارية المحدودة للأجهزة الذكية. لذلك ، من المهم التعامل مع هذه المشكلة عن طريق إيقاف تشغيل المستشعرات الزائدة التي تنقل البيانات المماثلة إلى الحوض وتفعيل عدد أقل من عقد الاستشعار في شبكة إنترنت الأشياء. هذا يقلل من القراءات المحسوسة الزائدة عن الحاجة ويقلل من الحمل الزائد للاتصالات ، وبالتالي ، إطالة عمر WSN.

في هذه الأطروحة ، تم اقتراح نهج جدولة موفر للطاقة باستخدام التجميع والتشجير في شبكات WSN. يتم تنفيذ النهج المقترح على مستويين: عقد الاستشعار والبوابة. تلتقط عُقد المستشعر البيانات وتضغطها باستخدام إما ترميز طول التشغيل المعدل (MRLE) أو ترميز هوفمان (HE). على مستوى البوابة ، تتم جدولة عُقد المستشعر بناءً على طرق التجميع لإنتاج أفضل جدول زمني لعقد المستشعر التي تظل نشطة في الفترة التالية. يتم تجميع عقد الاستشعار بناءً على ناقلات البيانات المستلمة في مجموعات وسيتم اختيار مستشعر واحد في كل مجموعة ليظل نشطاً في الفترة التالية.

تم إجراء نتائج المحاكاة (المحاكاة المخصصة في لغة Python) لإثبات فعالية التقنية المقترحة ، حيث توفر نتائج تنافسية مقارنة ببعض الأعمال الأخرى من حيث استهلاك الطاقة ونسبة المستشعر النشط ونسبة البيانات المرسله ونسبة البيانات المفقودة. أدى ترميز طول التشغيل التجميعي والمعدل (AgglomMRLE) إلى خفض البيانات المرسله والطاقة المستهلكة على التوالي من ٩٤,٣٪ إلى ٩٨,٨٪ ومن ٩٩,٥٪ إلى ٩٩,٨٪ مقارنةً بترشيح تردد البادئة (PFF). قلل ترميز Kmeans وطول التشغيل المعدل (KmeansMRLE) على التوالي البيانات المرسله والطاقة المستهلكة من ٧٥,٩٪ إلى ٧٨,٧٪ ومن ٩٥,٣٪ إلى ٩٨,١٪ مقارنة بـ PFF. خفض AgglomMRLE النسبة المئوية للبيانات المفقودة من ٥,٣٪ إلى ٢١,١٪ مقارنة بـ PFF.



جمهورية العراق
وزارة التعليم العالي و البحث العلمي
جامعة بابل كلية العلوم للبنات
قسم علوم الحاسوب

اسلوب جدولة كفوء الطاقة باستخدام العنقدة والترميز في شبكات الاستشعار اللاسلكية

رسالة مقدمة الى مجلس كلية العلوم للبنات في جامعة بابل كجزء من
متطلبات الحصول على درجة الماجستير في علوم الحاسبات

مقدمة من قبل

لينا وليد جواد

باشراف

الاستاذ الدكتور

علي كاظم ادريس السعدي