



# Multi-hop routing for cognitive radio sensor network

Mohamad Rida Mortada

## ► To cite this version:

Mohamad Rida Mortada. Multi-hop routing for cognitive radio sensor network. Networking and Internet Architecture [cs.NI]. ENSTA Bretagne - École nationale supérieure de techniques avancées Bretagne, 2021. English. NNT : 2021ENTA0010 . tel-03760771

**HAL Id: tel-03760771**

**<https://theses.hal.science/tel-03760771>**

Submitted on 25 Aug 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# Multi-hop Routing for Cognitive Radio Sensor Network

Mohamad Rida Mortada

## ► To cite this version:

Mohamad Rida Mortada. Multi-hop Routing for Cognitive Radio Sensor Network. Computer Science [cs]. Ecole nationale supérieure de techniques avancées Bretagne; American University of Culture and Education-Lebanon, 2021. English. tel-03527440

**HAL Id: tel-03527440**

**<https://tel.archives-ouvertes.fr/tel-03527440>**

Submitted on 15 Jan 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# THESE DE DOCTORAT DE

L'ECOLE NATIONALE SUPERIEURE  
DE TECHNIQUES AVANCEES BRETAGNE  
COMUE UNIVERSITE BRETAGNE LOIRE

ECOLE DOCTORALE N° 601  
*Mathématiques et Sciences et Technologies  
de l'Information et de la Communication*  
Spécialité : *Télécommunications, Informatique,*

Par

**Mohamad Rida MORTADA**

**Multi-hop Routing for Cognitive Radio Sensor Network**

Thèse présentée et soutenue à Brest, le 08 décembre 2021

Unité de recherche : Lab-STICC UMR CNRS 6285

## Membre du jury :

### Président :

Pr. Emanuel RADOI (UBO)

### Rapporteurs :

Pr. Yannick DEVILLE (Université Toulouse 3 Paul Sabatier - Toulouse)

Pr. M. Denis HAMAD (Université du Littoral Côte d'Opale)

### Examineurs :

Pr. Nadège THIRON-MOREAU (Université du Sud Toulon Var)

### Membres Invités :

Pr. Amer BAGHDADI (IMT Atlantique, CSI)

Dr. Christophe OSSWALD (ENSTA Bretagne, CSI)

Pr. Abdel-Ouahab BOUDRAA (Ecole Navale)

### Directeur :

Pr. Ali MANSOUR (ENSTA Bretagne)

### Co-Directeur :

Dr. Abbass NASSER (Doyen à AUCE)

### Encadrant :

Dr. Koffi-Clément YAO (UBO)



*“Whoever fears climbing mountains will live forever among the pits.”*

Tunisian poet Abu al-Qasim al-Shabi

”وَمَنْ يَتَّهِبُ صُعودَ الجِبَالِ .. يَعِشُ أَبَدَ الدهْرِ يَنْ الحُقَرِ.”  
الشاعر التونسي ابو القاسم الشابي

# *Abstract*

The Wireless Sensor Network (WSN) consists of a large number of tiny devices called nodes, and these nodes are generally limited in power and they are randomly deployed in a geographical area for monitoring purpose. Because of the large number of nodes in the WSN, their demand on the frequency resources becomes a real challenge due to the spectrum scarcity.

Cognitive Radio (CR) was introduced to enhance the spectral efficiency. The CR classifies users into the Primary User (PU) that holds a license over a spectrum bandwidth, and the Secondary User (SU), i.e., an opportunistic user. The integration of CR into WSN results in a cognitive radio sensor network (CRSN). In CRSN, the nodes behave as SUs. However, the adoption of CRSN may face several challenges. Indeed, network nodes may stop transmitting to avoid any harmful interference for PU. In addition, energy consumption constraint should be respected. Other features may be impacted by the adoption of CRSN, such as the clustering process and the in-network data aggregation, especially in a multi-hop routing based CRSN. In this manuscript, we tackle the challenges of the CRSN from different levels.

Firstly, clustering nodes enhances network efficiency. In practice, grouping the network nodes saves energy during data transmission. Thus, we investigate the optimal number of clusters in the network based on energy consumption during data transmission and spectrum sensing to extend network lifespan.

Then, we present LIBRO, a new uplink multi-hop routing protocol based on the geographical location information. LIBRO ensures delivery of rectified data packets in dense networks without knowledge of topology or path nodes.

Finally, we used LIBRO and the energy harvesting technique to extend the CRSN's lifespan. Thus, by solving an optimization problem jointing the collision probability, energy consumption, packet delivery delay and packet delivery ratio, the appropriate parameters are found.

**Keywords:** Wireless Sensor Network, Cognitive radio, Clustering, Network efficiency, Multi-hop routing protocol, optimization, Energy Harvesting.

# *Acknowledgements*

I would like to thank the members of the jury for accepting arbitration in my thesis. I especially thank the president Pr. Emanuel RADOI, the rapporteurs Pr. Yannick DEVILLE and Pr. M. Denis HAMAD and the examiner Pr. Nadège THIRON-MOREAU.

I'd want to convey my heartfelt gratitude and admiration to my director, M. Ali Mansour, Professor at Ensta-Bretagne; you've been a fantastic mentor to me. I'd want to express my gratitude for your support of my study and for helping me to advance as a research scientist. Your guidance on both research and my career has been invaluable. My heartfelt appreciation goes to M. Koffi-Clément Yao, Maitre de Conférence at UBO, for his unwavering support of my Ph.D. studies and associated research, as well as his patience and inspiration. Their advice was invaluable during the research and writing of this thesis. For my Ph.D. studies, I could not have asked for a finer advisor.

I would also like to thank to my PhD advisors in Lebanon, Dr. Abbass Nasser dean of the Faculty of Science at AUCE, for his great support and confidence in my ability to complete this work, as he was always ready and on call whenever I needed him, and he never hesitated to help me and meet me by all available means. Thanks to his constant follow-up to me in my work, pushing me to complete my tasks in the required time, and giving important advice, which led to the completion of my work, also, his technical and editorial advice was essential to the completion of this dissertation.

There is a person to whom I would like to thank a lot. Without him, all of this would not have started in the first place, Dr. Hassan Harb. You were like a gift from heaven. Without you, I would not have started my Ph.D. research, and I would not have met my PhD team in the first place. You were the main cause and the spark that started all of this. Thank you my friend.

I'd like to express my gratitude to my family in particular. I am indebted to my mother Fatima and father Tamer for all the sacrifices they made on my behalf. Your prayers for me have kept me going so far. I'd also like to thank all of my brothers, sister, and friends who have encouraged me to keep writing.

Finally, I would like to thank my wife, Batoul, Who supported me since the beginning by standing by my side and encouraging me to continue. Her trust in me and her belief in my ability to continue is an expression of her sincere love for me.

# Contents

<b>Abstract</b>	<b>ii</b>
<b>Acknowledgements</b>	<b>iii</b>
<b>List of Figures</b>	<b>vii</b>
<b>Abbreviations</b>	<b>ix</b>
<b>1 General Introduction</b>	<b>1</b>
1.1 WSN	1
1.1.1 Types of WSN (Periodic, Event Driven, On Demand)	2
1.1.2 WSN nodes different parts and limitations	2
1.1.3 Data gathering and challenging	3
1.2 Clustering in WSN	3
1.2.1 Centralized clustering	4
1.2.2 Decentralized clustering	4
1.3 Multi-hop Routing algorithms	5
1.3.1 Proactive routing protocols	6
1.3.2 Reactive (On demand) routing protocols	6
1.3.3 Data aggregation	6
1.4 CR Concept	7
1.4.1 Dynamic Access Spectrum	7
1.4.2 Spectrum access paradigm	8
1.4.2.1 Underlay Access	8
1.4.2.2 Overlay Access	9
1.4.2.3 Interweave Access	9
1.4.3 Spectrum Sensing techniques	10
1.5 Adoption of CR in WSN (CRSN)	10
1.5.1 Advantages	10
1.5.2 Limitation and challenges	10
1.6 Energy harvesting in CRSN	11
1.6.1 Concept	11
1.6.2 Sources and methods	12
1.6.3 Energy harvesting advantages	13
1.7 Conclusion and theses structure	13



<b>2</b>	<b>Optimal Clustering In Cognitive Wireless Sensor Network</b>	<b>15</b>
2.1	Related works and contributions . . . . .	15
2.2	CRSN's System Model . . . . .	19
2.2.1	Ring models and K-Means Clustering . . . . .	19
2.2.2	Multi-hops routing process . . . . .	21
2.2.3	In network data aggregation . . . . .	23
2.3	Problem Formulation . . . . .	24
2.3.1	SS energy consumption and lifespan . . . . .	25
2.3.2	Optimal number of clusters . . . . .	27
2.4	Network Latency and Collision Probability . . . . .	28
2.4.1	Collision Probability and Latency . . . . .	29
2.4.2	Maximal Consumed Energy . . . . .	32
2.5	Evaluation of our system . . . . .	33
2.6	Conclusion . . . . .	38
<b>3</b>	<b>A new routing algorithm for WSN (LIBRO)</b>	<b>40</b>
3.1	Introduction and related works . . . . .	40
3.1.1	Some multi-hop routing protocol in literature . . . . .	41
3.1.2	On demand geo-location routing algorithms . . . . .	42
3.1.3	TBF and DIR algorithm . . . . .	42
3.1.4	Presentation and contribution of LIBRO . . . . .	43
3.2	Location Information Based Routing Protocol (LIBRO). . . . .	43
3.2.1	LIBRO contributions . . . . .	45
3.2.2	Presentation of our protocol . . . . .	46
3.2.3	LIBRO Packet header and flowchart . . . . .	47
3.2.4	LIBRO energy consumption . . . . .	49
3.2.5	Packet loss probability . . . . .	51
3.2.6	System latency . . . . .	52
3.3	Evaluation of our protocol . . . . .	52
3.3.1	Energy consumption . . . . .	53
3.3.2	Packet loss and latency . . . . .	54
3.4	Conclusion . . . . .	55
<b>4</b>	<b>Routing Optimization In Energy Harvesting Cognitive Radio Sensor Network</b>	<b>56</b>
4.1	Introduction and Related works . . . . .	56
4.1.1	Related Work on Energy Harvesting . . . . .	57
4.1.2	Chapter's Contribution . . . . .	58
4.2	System Model . . . . .	59
4.2.1	CR Network Model . . . . .	59
4.2.2	CR-LIBRO routing model . . . . .	62
4.2.3	Energy harvesting routing model . . . . .	64
4.3	Problem Formulation . . . . .	64
4.3.1	Collision Probability . . . . .	64
4.3.2	Energy Consumption and harvesting . . . . .	65
4.3.3	Transmission packet latency . . . . .	66
4.3.4	Packet Delivery Ratio . . . . .	66

4.3.5	Loop-free condition . . . . .	67
4.3.6	Optimization problem formulation and resolution . . . . .	67
4.4	Numerical results. . . . .	69
4.4.1	Comparison with AODV and DIR . . . . .	69
4.4.2	Evaluation of optimization results . . . . .	73
4.5	Conclusion . . . . .	74
<b>5</b>	<b>General Conclusion</b>	<b>76</b>
5.1	Contributions . . . . .	77
5.2	Future works . . . . .	78
<b>6</b>	<b>List Of Publications</b>	<b>81</b>
<b>A</b>	<b>Appendix A</b>	<b>82</b>
A.1	Proof of Eq. (2.20) . . . . .	82
A.2	Proof of Eq. (2.21) . . . . .	84
A.3	Proof of Eq. (2.26) . . . . .	85
A.4	Proof of Eq.(3.10) . . . . .	86
A.5	Proof of Eq.(4.6) . . . . .	88
A.6	Proof of Eq.(3.13) . . . . .	88
A.7	Proof of Eq.(4.7) . . . . .	89
A.8	Proof of Eq.(4.10) . . . . .	90
A.9	Proof of Eq.(4.15) . . . . .	91
<b>B</b>	<b>Appendix B</b>	<b>92</b>
B.1	Simulation of double integral in equations Eq.(3.7) and Eq.(3.9) . . . . .	92
B.1.1	Java program: . . . . .	92
B.1.2	Results: . . . . .	93
	<b>Bibliography</b>	<b>94</b>

# List of Figures

1.1	WSN node internal architecture. . . . .	3
1.2	Multi-hop routing in WSN. . . . .	5
1.3	Data aggregation at CHs level. . . . .	7
1.4	SU spectrum access types. . . . .	9
1.5	Main grid-powered WSN. . . . .	12
2.1	Clustering Network with equal cluster size where $k = 3$ is the number of rings splitting a network of radius $R$ . . . . .	19
2.2	K-means applied on a circular CRSN uniformly distributed, the number of cluster was predefined. . . . .	22
2.3	Gathering data from CNs and CHs . . . . .	22
2.4	Data Routing between CHs located in different similar physical value zones . . . . .	23
2.5	Physical zones in CRSN . . . . .	24
2.6	Theoretical and simulation results of the lifespan vs number of rings in CRSN . . . . .	28
2.7	The PU activity modelled as a two-state Markov chain . . . . .	29
2.8	Comparison between the theoretical and the simulation result for the average number of relayed packets . . . . .	32
2.9	The effects of the ring number on the energy consumption per iteration on $e_{max}$ with different network densities . . . . .	33
2.10	The effects of the number of nodes on the energy consumption with different values of $k$ . . . . .	34
2.11	The effects of the number of nodes on Lifespan with respect to various values of $k$ . . . . .	34
2.12	The collision probability of CH with PU during its activity with and without IDA. . . . .	35
2.13	The duration of complete collecting data cycle from CN to the BS with and without IDA . . . . .	36
2.14	The collision probability of CH with PU during its activity with and without IDA with different rings . . . . .	36
2.15	The collision probability of CH with PU during its activity among all network rings . . . . .	37
2.16	Comparison of energy consumption between our system and the DSAC . . . . .	38
2.17	Comparison of data latency between our system and the DSAC . . . . .	38
2.18	Comparison of collision with the PU transmission between our system and the DSAC . . . . .	39
3.1	Distance Routing protocol an example . . . . .	43
3.2	LIBRO uplink and downlink data transmission. . . . .	45

3.3	$x$ and $y$ steps. . . . .	47
3.4	LIBRO packet header. . . . .	48
3.5	LIBRO flow chart. . . . .	48
3.6	The exact transmission range of $M$ which cover the whole next zone $Z_i$ . . . . .	49
3.7	The distance between a node $M$ in a previous zone, $Z_{i-1}$ , outer intersection of the next zone $Z_i$ with the $x$ axis. . . . .	50
3.8	Comparison table between LIBRO and DIR. . . . .	53
3.9	Lifespan comparison between LIBRO and DIR protocols . . . . .	53
3.10	Packet loss probability between source and destination. . . . .	54
3.11	Latency of transmission of a packet from source to destination. . . . .	54
4.1	Cognitive Radio Sensor Network. . . . .	61
4.2	(a) Control message through CCC, (b) Physical data message. . . . .	62
4.3	CR-LIBRO flow chart. . . . .	63
4.4	The number of collision between SU and PU occurred during the overall packet delivery period with different node number. . . . .	69
4.5	The number of collision between SU and PU occurred during the overall packet delivery period with different PU number. . . . .	70
4.6	Consumed energy by the overall network during the overall packet delivery period with different nodes number. . . . .	70
4.7	Consumed energy by the overall network during the overall packet delivery period with different PU number. . . . .	71
4.8	Delay between generating the data message by a target node, and receiving it by the BS, with different nodes number. . . . .	71
4.9	Delay between generating the data message by a target node, and receiving it by the BS, with different PU number. . . . .	72
4.10	Comparison of collision probability with different values of transmission range and the optimal value. . . . .	72
4.11	Comparison table between LIBRO, DIR and AODV. . . . .	73
4.12	Comparison of energy consumption with different values of transmission range, zone area and the optimal value. . . . .	74
4.13	Comparison of packet delivery delay with different values of transmission range, zone area and the optimal value. . . . .	74
4.14	Comparison of packet delivery ratio with different values of transmission range, zone area and the optimal value. . . . .	75
A.1	The forwarding packets coming from the physical ring $j$ by the CHs in spectrum ring $i$ . . . . .	83
A.2	Caption . . . . .	87
A.3	Packet transmission cycle . . . . .	90

# Abbreviations

<b>ADC</b>	Analog to Digital Converter
<b>AODV</b>	Ad-hoc On Demand Distance Vector
<b>BS</b>	Base Station
<b>CCC</b>	Common Control Channel
<b>CR</b>	Cognitive Radio
<b>CRSN</b>	Cognitive Radio Sensor Network
<b>CH</b>	Cluster Head
<b>CN</b>	Cluster Node
<b>CR-LIBRO</b>	Adopted LIBRO in CR
<b>DIR</b>	Distance Routing Protocol
<b>DSA</b>	Dynamic Access Spectrum
<b>DSAC</b>	Distributed Spectrum-Aware Clustering
<b>DSR</b>	Dynamic Source Routing
<b>EH</b>	Energy Harvesting
<b>FND</b>	First Node Dies
<b>GPS</b>	Global Positioning System
<b>IDA</b>	In network Data Aggregation
<b>LAR</b>	Location Aided Routing
<b>LEACH</b>	Low Energy Adaptive Clustering Hierarchy
<b>LIBRO</b>	Location Information Based Routing Protocol
<b>MCU</b>	Microcontroller
<b>MDP</b>	Markov Decision Process
<b>MHR</b>	Multi-Hop Routing
<b>MLDAS</b>	Minimum Latency Data Aggregation Scheduling
<b>OWC</b>	Optical Wireless Communications
<b>POMDP</b>	Partial Observable Markov Decision Process
<b>PU</b>	Primary User
<b>QoS</b>	Quality of Service
<b>RF</b>	Radio Frequency
<b>RSS</b>	Received Signal Strength

---

<b>SS</b>	Spectrum Sensing
<b>SU</b>	Secondary User
<b>TBF</b>	Trajectory-Based Forwarding
<b>TDMA</b>	Time Division Multiple Access
<b>WSN</b>	Wireless Sensor Network
<b>ZAP</b>	Zone Active Period
<b>ZRP</b>	Zone Routing Protocol

*For my friends...*

# Chapter 1

## General Introduction

It seems that the information revolution has a very big role in the progress and development of mankind at all the industrial, commercial, service and even cultural levels. The basic building block of information systems is the data that has become nowadays as the fuel that drives those systems to facilitate our lives.

The Wireless Sensor Networks (WSN) is a rich source of data [1]. These networks monitor the physical and environmental conditions such as temperature, humidity, atmospheric pressure, pollution rate, etc., and then collect the data and send it to the end user. Since these networks use wireless transmission to communicate among their components, they therefore need spectral resources to be able to deliver data from the geographical area to the user. The WSN uses wireless resources greatly [2], which makes the used channel congested, prompting researchers to find solutions to alleviate this problem.

The concept of Cognitive Radio (CR) comes to solve congestion on wireless spectrum resources [3], and use the spectrum resource in an efficient manner [4, 5]. CR aims to share the spectrum between Secondary Users (SUs) and Primary Users (PUs). SU can access the frequency channel only in the absence of a PU. Thus the SU must be aware of the status of the PU.

### 1.1 WSN

WSN contains a set of separated tiny devices, called nodes, usually deployed over a large geographical area. WSNs aim to collect physical data such as temperature, pressure, humidity *etc*, over a specific area, and forward them to intended users [6]. WSN nodes are battery powered devices with low energy Radio Frequency (RF) modules and processing



capabilities [7]. This node, basically, consists of one or multi physical sensor, wireless unit, power supply and micro-controller which controls and coordinates all of nodes equipment. Each node gathers local physical data, then transmit it to a Base Station (BS) in a single or multi-hop route [8].

### 1.1.1 Types of WSN (Periodic, Event Driven, On Demand)

WSN can function under three modes of operation [9]:

- Periodic mode: In this mode, the nodes in the WSN collect data from the geographic environment, and then send that information periodically to the BS.
- Event Driven mode: In this mode, the sensors do not send any data as long as the readings do not exceed a certain threshold. As soon as a physical value above the threshold is picked up by a sensor inside the network, this sensor starts sending data towards the BS.
- On Demand mode: In this mode, all sensors remain on standby, awaiting user commands. When the user requests data, the network controller broadcasts commands to collect data from the nodes, the nodes then send the collected data to the BS.

### 1.1.2 WSN nodes different parts and limitations

Fig. 1.1 depicts the internal architecture of a generic WSN node [7, 10]. As we can see, a single node is composed of a sensor unit, Analog to Digital Converter (ADC), Micro Controller Unit (MCU) and its memory, a radio transceiver and a power unit. The sensor unit reacts to physical signals from the surrounding environment, and converts them into a valuable electrical signal that reflects the measured physical values. The ADC converts the analog signals into digital ones. MCU is a low power processor with limited processing capability, this MCU is provided by a small memory which contains the program driving the node, with the capacity to stores a finite data array. The transceiver may be a short radio range unit which connect the node with a number of channels in order to communicate with other nodes or with the BS, as well as it receives the control commands coming from its neighbors or the BS. The power unit plays a vital role in providing all parts of the device with the necessary energy. In general, it may be difficult to replace the power unit when it runs out of power. This point must be taken into account when installing the operating system in the contract, as it must manage the energy consumption with high efficiency, and this is one of the most important challenges that inspire researchers in WSN.

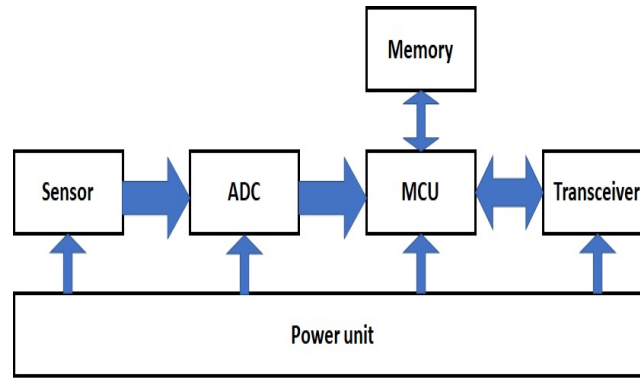


FIGURE 1.1: WSN node internal architecture.

### 1.1.3 Data gathering and challenging

Data gathering is the process by which the readings stored inside each sensor in the network are obtained, these readings reflect the status of the monitored area. Data is generally gathered via wireless transmission from sensor nodes to the BS.

Gathering a large amount of sensed data is very challenging in WSN. In fact, during data gathering, the WSN nodes relay their data toward the BS [11]. In a dense and large network, gathering data process faces two major problems:

- The short communication range of the sensor nodes, deployed sensors in a large geographical zone cannot communicate with each other directly, this makes it difficult to collect data from the entire network.
- In order to collect data from a sensors located far from the BS, the sensors near to the BS should relay data message generated by surrounding neighboring sensors, which causes that sensor to run out of power quickly.

As a result, clustering and multi-hop routing are offered as two potential strategies for circumventing the limitations mentioned above.

## 1.2 Clustering in WSN

Clustering is a process which divides the network nodes into groups. Clustering is performed according to pre-defined criteria, such as the distance separating the nodes, or on the basis of the convergence of the data collected from the surrounding environment. Each cluster must have a controller, called the Cluster Head (CH). The CH plays a key role in the cluster, as it collects data from the cluster nodes, then aggregates them into

one packet, and then sends it to the BS. Clustering of nodes in wireless sensor networks significantly improves network efficiency [12]. In fact, dividing the network into groups has a significant role in reducing the energy consumed in the data transmission process. It also reduces the need for wireless resources during data collection, as the transmission range of a cluster node is limited within a single cluster. Thus, clustering enhances data collection from the network by reducing network latency, as CHs can collect data from their cluster nodes simultaneously [12]. In addition, the process of collecting data from each cluster can contain similar information due to its geographical proximity. Therefore, the CH eliminates any redundancy in the incoming data, which is reflected positively in energy savings [13].

### 1.2.1 Centralized clustering

A centralized clustering is a type of clustering where the division of the network into clusters is the responsibility of the base. In this type of clustering, the nodes in general, periodically transmit information about their geographic location and residual energy to the BS. The BS, in its turn, performs the necessary processing to find the best nodes to play the role of the CHs in the network. The BS then broadcasts the identities of the nodes that will be the CHs during the current round. In the end, each node receives its identity from the BS message, becomes a CH, and then broadcasts its identity in order to inform nearby nodes of the presence of a nearby CH [14]. Nodes other than CHs choose their clusters by choosing their closest CH to transmit their collected data. The advantage of central clustering is in its simplicity, making the BS the center of clustering decision-making, as a result, individual nodes are not burdened with the clustering phase processing load. The disadvantage of this technology is that, in general, it consumes a lot of nodes energy when collecting information about geographic location and remaining energy. It also consumes a lot of spectral resources in the network.

### 1.2.2 Decentralized clustering

Decentralized clustering is the division of a network into clusters without any need for a central controller [15]. This type of clustering allows the cluster and CH to be identified locally and automatically without the need for external intervention, allowing for network scalability and energy savings because there is no need to relay geographical position data and remaining energy to the BS periodically, as there is with centralised clustering. It also saves wireless resources in the network. One of the examples that should be mentioned in decentralized clustering is the LEACH (Low Energy Adaptive Clustering Hierarchy) algorithm [16], since this algorithm is fully distributed over the

nodes. The working mechanism is as follows, each node chooses a uniformly distributed random number between 0 and 1, if the chosen number is less than a certain threshold, the node considers itself a CH and then broadcasts its identity to the local nodes. Nodes that choose a number higher than the threshold, are considered normal nodes, and then determine their belonging to the cluster based on Received Signal Strength (RSS) coming from the CH. These nodes then send the data to their appropriate CH, which in its turn forwards it toward the BS. This mechanism is repeated periodically, allowing other nodes to play the role of the CHs, which distributes the load fairly across all nodes in the network. It is difficult to implement this algorithm in large scale networks without a Multi-hop routing mechanism, due to the inability of a far CH to send data directly to the BS because of energy constraints in WSN.

### 1.3 Multi-hop Routing algorithms

Since nodes within the network have a limited transmission range, also a source and a BS may not share a common available channel to communicate directly, they cannot send data in a single hop manner to the BS, especially in large scale WSN [17]. Therefore, Multi-Hop Routing (MHR) is a solution for transmitting data from source node to the BS in multi-hop manner as shown in Fig. 1.2. Multi-hop routing reduces energy consumption in the network, which extends the network lifetime, as well as, it reduces the reservation of a lot of wireless resources as well as the inter-collision between two transmissions, due to the adoption of a short-range transmission.

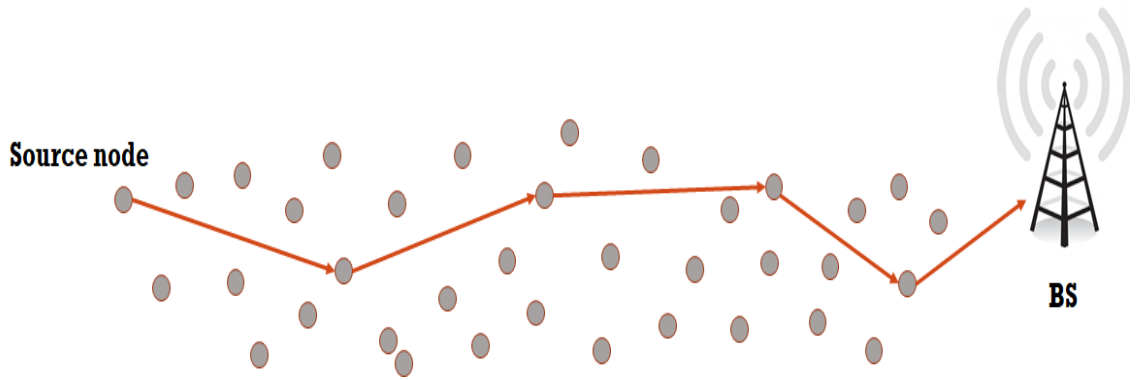


FIGURE 1.2: Multi-hop routing in WSN.

### 1.3.1 Proactive routing protocols

In this approach [18], each node within the network maintains a path to each assumed destination. Since routing data is generally stored in a table within each individual node, the routing table consists of several entries, each of them contains the path to a specific destination. Routing data is regularly updated by sending control packets such as a discovery route through the network. This technique is effective in networks where immediate transmission of data without significant delays is required. However, regular transmission of control packets consumes a lot of power and memory space in the network nodes, especially since nodes can hold/update information of a route that will never be used.

### 1.3.2 Reactive (On demand) routing protocols

In order to encounter the proactive routing due to limited network nodes resources such as power, small memory space, and limited processing, the reactive routing [19] has been proposed for WSN. Whereby the route is formed when it is needed; When a node needs to send data to a specific destination, it first performs the route discovery procedure, then chooses the appropriate route consisting of a certain number of nodes between the source and destination, and then the data transmission process begins. Once the data transmission process is completed, this route is deleted from the cache memory of its component nodes.

### 1.3.3 Data aggregation

Data aggregation in clustered WSN is a very important process in reducing the amount of data sent from the CHs to the BS. This process has a significant impact on saving energy consumption and reducing the latency in sending data from the source to the BS. The process of data aggregation starts inside network clusters, where each node sends its data to its appropriate CH, and then the CH aggregates that collected data according to a pre-defined aggregation function such as the sum, maximum value, average value or minimum value [20]. Then the CH sends all the collected data in one message to the BS. The process of sending data from the CH to the BS can be in MHR manner among the other CHs towards the BS, and thus the process of data aggregation can be performed, so that two different CHs, while sending their data to the BS, can overlap on their way at a particular CH as shown in Fig. 1.3, CH0 and CH1 transmit their data packet to the CH2. In its turn, CH2 by duplicate-suppression function it eliminates the repetition and converting the two messages into one message before sending it back to

the CH3. This process is called Data-centric Protocol or In-network Data Aggregation (IDA) [21–23].

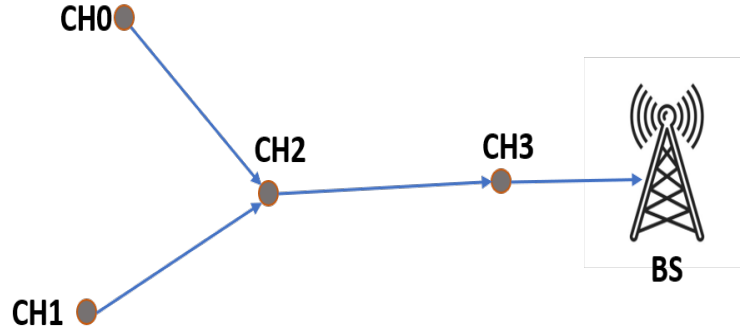


FIGURE 1.3: Data aggregation at CHs level.

## 1.4 CR Concept

Because of recent developments toward 5G service, as well as significant advancements in the Internet of Things (IoT) and a wide range of modern wireless applications, the wireless communication community is confronted with an enormous problem of how to adapt limited frequency resources to the growing demand for wireless communication services [24]. As result of this development in the process of wireless communication and multimedia services, the radio frequency spectrum is becoming increasingly congested, and it becomes highly necessary to look for alternative frequency resources.

### 1.4.1 Dynamic Access Spectrum

A promising proposed strategy for addressing spectrum scarcity in the radio frequency (RF) spectrum, Dynamic Spectrum Assignment (DSA), has emerged in recent years [25]. However, the vast majority of communication systems employ predefined communication parameters and assign a specific frequency, resulting in systems designed to work only on predetermined frequency bands. In light of recent studies indicating that the assigned frequency bands account for anywhere between 15% and 85% percent of the total available wireless spectrum, this flat distribution results in low spectrum use efficiency [26]. Because of this low efficiency and spectrum scarcity, several advancements and developments in wireless communication has been proposed. The wireless community developed DSA as a solution to achieve prudent and efficient use of the spectrum [27, 28], which results in the acquisition of additional frequency resources. In this regard, CR has

been suggested as a contender for DSA performance by taking advantage of available frequency bands in order to maximise efficiency. Another alternative to radio frequencies, there are various studies that use Optical Wireless Communications (OWC) to transmit data. OWC is one of the proposed solutions to the spectrum shortage problem. Indeed, OWC communicates between two sites via optical link rather than radio frequency (RF) channel. OWC enables high data rates while consuming minimal power [29]. On the other hand, OWC faces numerous constraints, including a small transmission distance and a significant susceptibility to environmental conditions such as rain, snow, fog, and aerosol pollutants [30].

### 1.4.2 Spectrum access paradigm

With the increasing demand on the wireless communication technologies, the available frequency resources become scarce. However, a large portion of the licensed spectrum is severely under utilization. Initially, CR was proposed by Mitola [31]. This system is characterized by its ability to perceive the spectrum and take advantage of the opportunity to use unused channels. In addition, wireless devices that use this system are characterized by the ability to automatically adapt to changes in the spectrum, as they can adjust communication parameters such as modulation, channel coding and frequency range *etc.* Due to its ability to distinguish its spectral surroundings, CR divides users into two categories: the PU, this user is licensed, he can use his licensed channels absolutely freely whenever and for as long as he wants. The SU, this user can take advantage of the opportunity to use any licensed channel provided that during his use of it, it does not cause any interference or impact on the Quality of Service (QoS) for the PU.

For optimal use of the spectrum, the PU and the SU must share the spectrum in an efficient manner. Therefore, as shown in Fig. 1.4 three models for accessing the spectrum are proposed:

- Underlay Access
- Overlay Access
- Interweave Access

#### 1.4.2.1 Underlay Access

In this model, a SU can simultaneously access a channel used by a PU [32]. Thus, the SU must respect that the maximum power of its transmission signal does not exceed

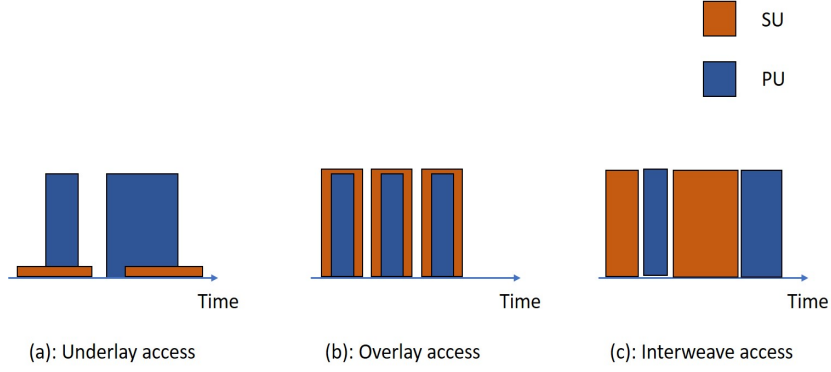


FIGURE 1.4: SU spectrum access types.

the limit specified by the Federal Communication Commission (FCC) under the name of the interference temperature [33]. To be able to accomplish this the SU can spread its signals over a very wide frequency range to ensure that the transmitted power is less than the power of any transmitting PU. This type of transmission can be used in short ranges due to the limited transmitting power. The advantage of this technique is the ability of the SU to transmit simultaneously with the PU. The main challenge is in the techniques of determining the power of the interference signal arriving at the receiver in the PU and the possibility of ignoring it [34].

#### 1.4.2.2 Overlay Access

This model differs from the previous one, where the SU can send in full power simultaneously with the PU, provided that the transmission of the SU does not affect the QoS for the PU. This scenario is based on the full cooperation between the PU and the SU, so that the PU informs the SU of its data sequence, which enables the SU to perform the relaying between two PUs. While relaying the PUs, the SU sends its own data through network coding [35].

#### 1.4.2.3 Interweave Access

This model is the most attractive of the three access models of the last decade [36]. In this model, the SU accesses the spectrum in an opportunistic manner, whereby the SU can use a channel when the PU is absent, where they can transmit at full capacity. In order to avoid interference with the PU, the SU must vacate the channel as soon as the PU arrives, and then search for an inactive channel that is not used by the PU.



### 1.4.3 Spectrum Sensing techniques

Spectrum Sensing (SS) [36] is the key function in the CR interweave access model. Each time a SU accesses a channel, it must sense the spectrum, and then decides whether or not that channel is currently being used by a PU. The SS algorithm must be robust and fast, which gives correct results in a relatively short time. If a SU decides that an active channel is idle, this will result in a collision with the PU transmission. On the other hand, if a SU incorrectly determines that a channel is active while it is inactive, the opportunity to use that channel will be lost.

## 1.5 Adoption of CR in WSN (CRSN)

CR-based communication has been proposed to supply extra spectral resources to the WSN [5]. All WSN nodes (including the end-nodes, CHs, and BS) operate as SUs in CR-based WSN (CRSN). To avoid any interference, they can only access the PU channel when the latter is not present.

### 1.5.1 Advantages

The amount of data transmitted can be very large, especially in large and dense WSN. Thus, networks face difficulty in transmitting data at the data aggregation phase, especially if the spectrum is crowded with other users such as PUs. Hence, WSN data transmission collide with other PUs transmission, which leads to the data packet loss problem. The local transmission of a sensor node in the network may also collide with the transmission of another sensor node in the same network. By performing SS, the CRSN can avoid these interferences during the transmission process, and it also searches for other unused channels, which speeds up the data collection process, and optimizes the use of the spectrum [37]. In addition, because of its ability to avoid collisions during transmission, this CRSN can reduce the amount of re-transmitted data needed to ensure that the message reaches its destination. Thus, energy consumption is reduced in those networks, which increases the lifespan of the network.

### 1.5.2 Limitation and challenges

CRSN networks provide a smart solution to solve the problem of spectrum scarcity, but they face a number of challenges while applying them to WSN [4]. In some cases, they may be imposed on the sensor nodes in the network by stopping repeatedly during

transmission, and vacating the channels used by them to make place for the licensed PUs to use those channels freely without affecting their QoS. This frequent downtime can be an obstacle in some applications where transferring data at a certain time limit is critical, or in data-streaming applications. In some other applications, and to ensure good network productivity, priority may be given to certain nodes over other nodes, and here another challenge arises, which is the ability of the CRSN to properly manage those priorities. Another challenge lies in the optimal selection of CR algorithms that consume the least amount of energy in CRSN networks, due to the limited energy resources in those networks, which affects the overall lifespan of the network. There are also cases where several CRSNs are geographically overlapping and use the same spectrum band, which brings the networks into competition with each other for radio resources, which also creates new challenges in CRSNs.

## 1.6 Energy harvesting in CRSN

One of the most important challenges facing CRSNs is to rationalize energy consumption in an optimal way in order to extend the lifespan of the network for the longest possible period [38]. In battery-powered networks, energy consumption is a potentially critical process because of the difficulty of replacing batteries in the network nodes due to the large number of nodes in such network, or because the network is located over a large scale geographical area and may also be rugged or located in a harsh area [39]. The network may contain a big number of batteries that may contain toxic materials that cause danger to the environment and the animals near that network [40]. In this case, the lifespan of the network is directly related to the ability of the batteries to supply it with energy.

In the main grid-powered networks (Fig. 1.5), nodes derive their energy from the electrical grid. Thus, this type of network is limited in the number of nodes and therefore it is difficult to extend these networks. In addition, it is not used in mobile networks due to the impossibility of powering them. These networks may be fed from an energy source that depends on fossil fuels, which increases the carbon footprints in the environment. Therefore, there was a need to search for renewable and environmentally friendly energy sources.

### 1.6.1 Concept

Energy Harvesting (EH) in CRSN networks needs to provide the nodes with a harvesting tool in order to extract energy from the environment and use it in the network's functions

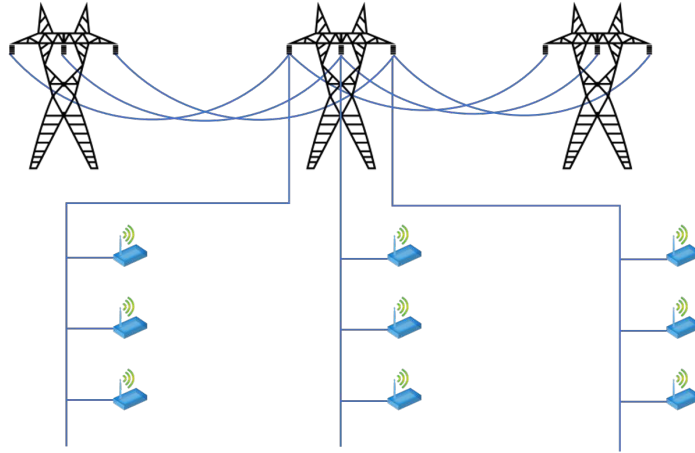


FIGURE 1.5: Main grid-powered WSN.

such as spectrum sensing, transmitting, receiving and switching to an idle channel [41]. There are two basic models in CRSN that support energy harvesting [42]: The first model is harvest-use, where the nodes harvest energy and then use it directly so that the energy harvested must be greater than the energy consumed in the same phase. The second model is harvest-storage-use. In this model, the nodes may be equipped with super-capacitors [43] which are less polluting to the environment than batteries. In this model, a large amount of energy can be stored for future use in multiple phases. The process of harvesting and using can be in a simultaneous manner or in a time-switching way.

### 1.6.2 Sources and methods

EH is applied in a way that allows devices to extract energy from the surrounding environment easily, from free, available and renewable energy sources. Solar, wind, biomass, tidal, hydro and even the energy contained in the radio frequency signal [44] are among the most important available energy sources that can be harvested. Harvesting energy from radio frequencies is an interesting topic, especially since the devices can harvest energy from the same received signal, so that the harvester splits the signal into two parts, a section that goes to store energy for use, and a section that carries data that goes for processing. It is similar to the way passive RFID tags work, where the RFID tag uses part of the energy of the received signal to resend its messages to the RFID reader [45]. In a CRSN, a device can transmit data on an idle spectrum and simultaneously harvest energy from an occupied one [46]. One of the advantages of radio frequencies is that it carries data and energy in the same signal [47]. Therefore, the BS in the CRSN network can send data at a high power, in order to allow the nodes to benefit from the energy of the radio signal while receiving data.

### 1.6.3 Energy harvesting advantages

One of the most important advantages of energy harvesting is liberating the network from energy constraints [48]. The CRSN network, which consumes less energy than it harvests, can exist permanently. The reliance of the CRSN network on harvesting energy from clean and renewable sources has a positive effect in reducing pollution, and thus these networks are environmentally friendly. By liberating the network from energy constraints, it allows the use of more diverse and effective algorithms to ensure that information is delivered to its destination and increase the reliability of SS process. It also allows the nodes to perform more complex data processing operations, as some data need complex sensors that may consume more energy. The challenge remains in the optimal selection of algorithms used in all network activities in order to increase its reliability and extend its lifespan as long as possible.

## 1.7 Conclusion and theses structure

In this chapter, we have defined the WSNs and their various uses in many fields, and we have classified them according to the way they use and collect data; Periodic, event driven and on demand data gathering. We also explained the basic parts of the sensor nodes, which are the basic building blocks of the WSN. These networks suffer from several limitations, the most prominent of which are listed. After that, we explained some of the data collection methods and the different challenges they face. We later defined the clustering in those networks and their positive impact in reducing energy consumption and accelerating the process of data collection, as it reduces the need for a large number of radio resources. We have divided clustering into two categories: central clustering and decentralized clustering, and presented the characteristics of each, their advantages and disadvantages. Then, we highlighted the urgent need for methods to ensure that data is successfully transmitted from the source to the final destination. We talked about the multi-hop routing algorithms and divided them into two types: the proactive and the reactive routing protocols. In addition, we explained the data aggregation processes and their importance in reducing the amount of data transmitted and thus saving more energy to extend the network lifespan.

CR concept has been also presented. CR can be applied in the WSN to overcome the spectrum scarcity resulted from the huge deployment and the large application of the WSN. We explained the dynamic access spectrum and its most prominent models such as: underlay access, overlay access and interwave access, as well as some of their features and limitations of each paradigm. In addition, we explained the importance of SS and its importance in determining the spectrum channel states, and we also showed the

importance of the sensing results being reliable because of their effective impact on the optimal use of the spectrum.

In order to energize the system with perpetual energy, energy harvesting technology has been introduced in these networks. In the last section of this chapter, we present energy harvesting techniques and their benefits in extending the life of the network and allowing it to use various algorithms that serve its vital functions. We have mentioned some of the most prominent sources of energy that can be harvested, especially the RF energy, its characteristics and advantages, and finally we mention the challenges facing energy harvesting.

The rest of this manuscript is divided as follow: In the chapter 2, we present a novel clustering optimization method based on the energy consumed during data transmission and SS phase in CRSNs. Our work includes the IDA technique which reduces the data latency and the energy consumption in our model. In chapter 3, we present a new routing protocol based on the geo-location information called LIBRO, such protocol allows the mobility of nodes inside the network. In LIBRO, each single node does not need any information about its neighbor node geo-locations or identities (IDs). Our protocol outperforms the state-of-the-art Distance Routing Protocol (DIR) in term of latency and energy consumption. In chapter 4, we present CR-LIBRO: the incorporation of our LIBRO protocol in the CRSN, we formulate a multi-objective optimization problem which joint the collision probability, energy consumption, data latency and throughput. A solution of this optimization problem is presented and our model is compared with DIR and the well known Ad hoc On-demand Distance Vector (AODV) routing protocol. Our system outperforms DIR and AODV in term of energy consumption, packet delivery delay and collision probability with PUs. Finally, we conclude our work and present new challenges with future work in the last chapter.

## Chapter 2

# Optimal Clustering In Cognitive Wireless Sensor Network

When a PU channel becomes inactive, the nodes in CRSN can access it. As a result, the nodes are assumed to have a SS module to track PU activity. In this chapter, we look at a clustered CRSN with an Ad-Hoc topology and IDA capability, where the SS operations are limited to CHs, which also gather data and send it toward a central BS. When the number of clusters in a CRSN grows, the energy consumed by data transmission decreases, while the total energy consumed by SS increases, because more CHs must perform SS before transmitting.

The impact of IDA on the performance of the CRSN is examined in this chapter. After determining the optimum number of clusters, a research aimed at extending the network lifetime is developed, taking into account SS needs, the IDA impact, and the energy consumed by both SS and transmission. In addition, the network delay and collision rate between primary and secondary transmissions are theoretically calculated. The effectiveness of IDA in extending network lifetime and reducing collision probability and network latency is supported by simulations.

### 2.1 Related works and contributions

Nowadays, CR is being used in WSN to address the issue of channel scarcity [49–51]. By sharing the frequency range between PU and SU, CRSN achieves excellent spectral efficiency. As long as a channel licensed to a PU remains idle, a sensor node in CRSN may access the spectrum on an opportunistic basis. The SU sensor node must sense the channel in order to keep track of the PU's state in order to prevent any unwanted interference. The SU sensor node should instantly leave the channel once the PU starts

its operations. As a result, the SU sensor nodes must be equipped with a SS unit [52–54]. Even while CRSNs improve spectrum efficiency by using unused or underutilised spectrum bandwidths, the increased efficiency comes at the expense of increased power consumption owing to the conducted sensing activities [55, 56].

Clustering, Multi-Hop Routing (MHR), and data aggregation are effective methods for reducing transmission energy while maintaining transmission quality and reliability [57–61]. As each node transmits its data through the CH to the central BS, they decrease the transmission radius of the nodes. Each node in a CRSN may conduct a variety of physical sensing tasks, including temperature, humidity, air pressure, and so on, before transmitting its data to its assigned CH. By expanding the number of clusters in the CRSN, the distance between nodes and their CHs may be reduced. Furthermore, clustering reduces the redundancy of the nodes' data because the CH can use an In-Network Data Aggregation (IDA) procedure (based on the mean, max, or min of the received data [62]) on the nodes and only transmit aggregated data to the next nearest CH toward the BS, which is the final destination using MHR.

Clustering and MHR in WSN and CRSN have been the main subjects of a number of studies. The authors of [63] suggest a protocol to help in cluster setup and data transfer. Their simulations demonstrated the efficacy of the suggested method, which increases network lifespan while decreasing energy usage. As a result, the authors of [64] suggest a clustering routing method that prevents low-energy nodes from becoming CH. Furthermore, residual energy is considered to be an energy-control element that helps balance energy consumption. The authors of [65] present an adaptive clustering routing method for WSNs to address the scalability issue in the case of large-scale WSNs. This method splits the network nodes into clusters before choosing one node from each cluster to serve as the CH. MHR routes are determined using the algorithm's energy balancing concept, which takes into account residual energy, load, and network size. The authors of [66] utilise the Gustafson-Kessel algorithm, which optimises the number of clusters to minimise energy consumption depending on cluster form and volume, initial clustering method configuration, data item distribution, and cluster number. The authors of [67] demonstrate that the Optimal Number of Cluster Heads (ONCH) method over LEACH (The Low Energy Adaptive Clustering Hierarchy Algorithm) is superior to LEACH without ONCH in terms of energy usage and sensor network longevity. The authors of [68] suggest a novel CRSN clustering method in which a pair of nodes that are close to each other and share the same channel may be linked together. During a single connection period, one of these two nodes should be put to sleep, while the other should remain active. Following that, at each interval, they switch states. Only awake nodes execute the SS and vote for the CH, which increases energy efficiency. To enhance their groupwise

spectrum-aware clustering method, the authors of [69] and [70] use the optimum cluster number. The authors of [69] propose a distributed spectrum-aware clustering method to make spectrum-aware clustering more reliable and practicable, whereas the authors of [70] provide an optimum cluster number to build up their energy aware cluster based routing protocol. The optimum number of clusters is calculated based on the energy spent in intra-cluster communication during a period of data collecting in each cluster, as well as the energy consumed when sending data in inter-cluster from each CH to the BS in one hop transmission. Using MHR [58], the number of relays increases with the hops between sources and destinations [71].

In our work, CH plays the role of a relaying node. As a result, increasing the number of clusters and CHs in the network is required to decrease the transmission range. On the one hand, since the radio transmission distance is shorter in MHR-based networks, the energy usage reduces. Due to the limited transmission range, the risk of colliding with the PU may be minimised. On the other hand, the latency and the error probability may increase [72]. However, in CRSN, another component, the SS, should be included. As a result, the optimum number of clusters in a CRSN should be determined not just by reducing energy consumption, but also by preserving the SS's goal performance.

Data aggregation, on the other hand, is one of the most used wireless network compression methods. Due to concurrent data collection, this method saves energy while collecting data and speeds up the process (each CH collects data independently and simultaneously with other CHs in the network). Several studies, including [73–76], address data aggregation for CRSN. Authors of [73] combine energy adaptive methods and information correlation for a multi-hop CRSN topology in an information theoretical capacity maximisation framework. In their simulations, they utilise a relatively tiny network (3 sensors per hop and 3 hops to the sink), and they find that an energy adaptive mechanism beats a non-adaptive mechanism. They also came to the conclusion that data aggregation networks that utilise information correlation outperform the others. Works of [74–76] deal with the Minimum Latency Data Aggregation Scheduling (MLDAS). In [74], a dense CRSN is considered. The authors postulate that SUs have asymmetric communication connections and that a SU-SU collision may occur. SU should operate in a competitive mode to minimise the number of collisions and to better share spectrum resources. In order to explore the MLDAS issue, two realistic distributed methods are suggested based on the unit disc graph interference model and the physical interference model. The authors of [75] concentrate on a probabilistic network model, which means that connections between nodes are not always guaranteed. A Regular Wireless Network with regular users and an Auxiliary Wireless Network with auxiliary users are the two networks that are examined. Any user may transmit data across the regular wireless



network spectrum on an equal chance with other users, or it can operate on auxiliary wireless network opportunistically if and only if the auxiliary wireless network spectrum is not occupied by any auxiliary user. In [76], authors build an aggregate tree and a conflict-free schedule at the same time, with no predefined structure. Multiple-channel access is added to MLDAS. Furthermore, since a scheduled node cannot participate in the aggregation process, a novel aggregation mode dubbed Data Aggregation Scheduling in the Dark is suggested to take use of scheduled node spectrum opportunities.

In CRSN nodes, energy is required for physical sensing, clustering election, data transfer, and other tasks. In addition to these activities, CRSN requires energy to perform SS. As a result of its unavoidable function, the CRSN must perform an energy-intensive task. Our research focuses on two possible energy consumption sources: SS and data transmission. In a conventional WSN [77], data transmission is the most energy-intensive activity, so other sources of energy consumption are overlooked. The SS in our model is solely CH's responsibility. Before collecting data from the cluster nodes and prior CHs, CH should run SS to avoid interfering with the PU. A small number of clusters results in significant energy consumption during the data transmission phase, as previously stated, due to the distance between CHs. A large number of clusters, on the other hand, causes significant energy loss during the SS phase. The CHs also carry out IDA in order to reduce data redundancy. The amount of data sent as well as the time it takes to send it will be affected. As a result, data aggregation and data delivery latency have an impact on the collision rate between PU and SU transmissions.

In this chapter, we look at how to optimise the CHs number within a CRSN given a variety of parameters, including the IDA, the energy consumed by physical sensing, the energy consumed during transmission, the SS performance that meets the predefined protection level for the PU against CRSN node interference, and the high spectrum efficiency that these nodes can exploit. The role of SS in determining the optimal number of clusters in a CRSN was not explicitly addressed in previous research works. Because the CRSN serves as a backup network, it must adhere to the SS requirements. The major contributions of this study is that, first, finding the optimal number of clusters of CRSN that takes into consideration several parameters, such as: the energy efficiency related to the transmission and the sensing operations, the detection and false alarm probabilities of clustered CRSN. Then, we obtain the average number of relayed packets by each CH in a closed form, thus we calculate the collision probability with the PU transmission. The latency of our energy-optimal clustered CRSN are derived in closed forms, taken into consideration the data redundancy reduction thanks to the

IDA. Finally, the numerical Results are conducted to show the efficiency of choosing the optimal number of clusters, and to figure out the related collision rate and the latency.

## 2.2 CRSN's System Model

Let us consider a CRSN that is operating on an opportunistic basis over a licensed channel that is reserved for a PU. We assume a deterministic network model with stable links between nodes for the sake of simplicity, and we also consider symmetrical communication links between SUs, the antenna gain at transmitter and receiver is 1. The CRSN contains  $N$  homogeneous sensor nodes [78] distributed over a large circular geographical zone [79] with a radius  $R$ . In this zone, all physical records should be transmitted toward a BS located at the center [80]. The CRSN is assumed operating on a band allocated to one PU, i.e. the PU activity throughout all the CRSN is homogeneous. Such an assumption is suitable to the realistic scenarios where the PU coverage is large and the whole CRSN falls into this coverage, e.g. the Global System Mobile (GSM) network [81]. Another case study is the Radio FM and TV white space, where the same frequency may be used/unused by the same PU within a whole country [82].

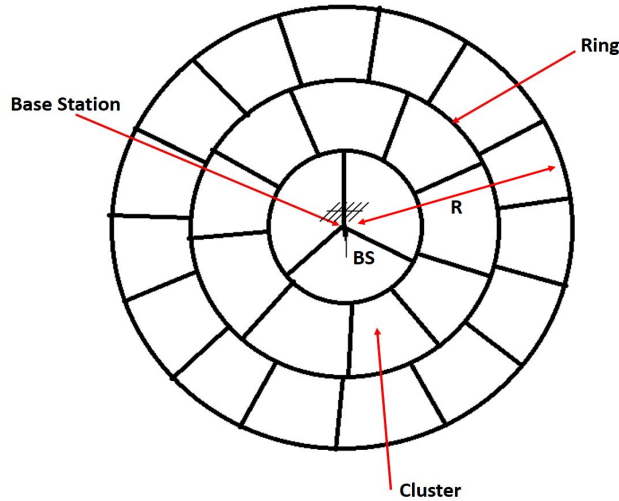


FIGURE 2.1: Clustering Network with equal cluster size where  $k = 3$  is the number of rings splitting a network of radius  $R$ .

### 2.2.1 Ring models and K-Means Clustering

As shown in Fig. 2.1, the network is divided into a set of  $k$  rings, referred to as SS rings (Spectrum Sensing rings). All ring widths are equals to  $\frac{R}{k}$ , and each ring is divided into

Notations List	
Notation	Definition
$N$	Total node numbers in the network
$M$	Number of CHs in the network
$R$	Geographical zone radius
$k$	Number of rings in network
$A_i$	Area of ring $i$
$m_i$	Number of cluster inside ring $i$
$M$	Total number of clusters in the network
$k_p$	Number of physical ring in the network
$p$	Number of distinct packet received by a CH
$p(H_0)$	Probability of the PU to be idle
$p(H_1)$	Probability of the PU to be active
$p_{fa}$	Probability of false alarm
$p_d$	Probability of miss-detection
$p_t$	Probability of transmission of a CH
$N_{p_i}$	Average number of relayed packets by CHs in the ring $i$
$e_r$	Energy density at the receiver antenna
$e_s$	Energy consumed during SS process
$e_0$	Energy consumed by the node in the running mode
$e_i$	Average of energy consumed by a node in ring $i$
$L$	Lifespan of the network
$e_{int}$	Initial stored energy of the node
$e_{max}$	Maximum energy consumed by a node within a cycle
$D$	Density of the nodes in the network
$K_{opt}$	Optimal number of rings
$e_{opt}$	Optimal consumed energy
$P_c$	Collision probability with PU during overall data gathering
$P_{c_i}$	Probability of collision with PU in the ring $i$
$\tau_s$	Time slot of the PU activity Markov model
$T_{a_i}$	Average time of transmission activity of CH in ring $i$
$T_{CN}$	Period of data packet transmission from one single CN to its CH
$\tau_{T_i}$	Average time of collecting data of a CH in ring $i$ from one of its previous CHs
$V_i$	Average number of previous CHs in the ring $i + 1$ of a CH in the ring $i$
$T_P$	Period of a single CH packet transmission
$mp_j$	Number of physical zones in a physical ring $j$

TABLE 2.1: The notations.

several equal-sized clusters called SS clusters. If the first ring (near the BS) contains three clusters, the area of each cluster is equal to:  $\frac{\pi}{3} \left(\frac{R}{k}\right)^2$ . The area of the  $i_{th}$  ring can be evaluated as follows:

$$\begin{aligned} A_i &= \pi \left(i \frac{R}{k}\right)^2 - \pi \left((i-1) \frac{R}{k}\right)^2 \\ &= \pi(2i-1) \left(\frac{R}{k}\right)^2 \end{aligned} \quad (2.1)$$

where  $i \in [1, k]$  is the index of the ring. Hence, the  $i_{th}$  ring contains a total number  $m_i$  of clusters equals to:

$$m_i = \frac{\pi \left(\frac{R}{k}\right)^2 (2i-1)}{\frac{\pi}{3} \left(\frac{R}{k}\right)^2} = 3(2i-1) \quad (2.2)$$

Assuming that the closest ring to BS ( $i = 1$ ) contains 3 clusters.

Using Eq. (2.2), we can find the total number of clusters  $M$  in the overall network:

$$M = \sum_{i=1}^k 3(2i-1) = 3k^2 \quad (2.3)$$

By definition, K-Mean is a well-known algorithm in network clustering that splits a network of  $N$  nodes located in a  $n$  dimension space into a  $K$  cluster ( $K$  clusters differ from  $k$  rings in our system) based on the Euclidean distance between clusters [83]. In WSNs, the goal of the K-Mean is to reduce the within-cluster sum of square distances. Thus, the topology of our adopted system model is inspired from the behavior of the K-means clustering algorithm when applied to a circular WSN. Fig. 2.2 gives an example of the K-mean clustering, in which the topology may be approximated to our model with 3 rings. As shown in this figure, the ring around the origin (where the BS is located) contains three clusters, similarly to our system model.

Our network is considered as an homogeneous network. However, we can as well extend our work to a heterogeneous network with  $A$  nodes where only  $N$  nodes,  $N \leq A$ , can be elected as CHs, because they have the required capability to fulfill the *CH* role.

### 2.2.2 Multi-hops routing process

In our model, multi-hops transmission system is adopted among the CHs. Thus, the message sent from a far CH to BS should pass by the CHs in between.

According to our model, CRSN carries out an IDA technique [84]; where CH gathers the data of the nodes inside its cluster and aggregates the data collected from farther

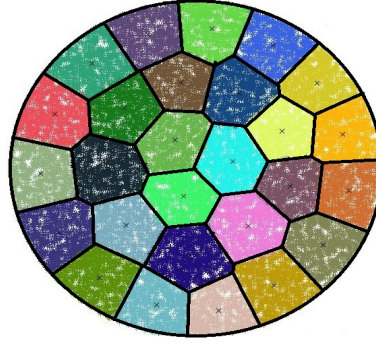


FIGURE 2.2: K-means applied on a circular CRSN uniformly distributed, the number of cluster was predefined.

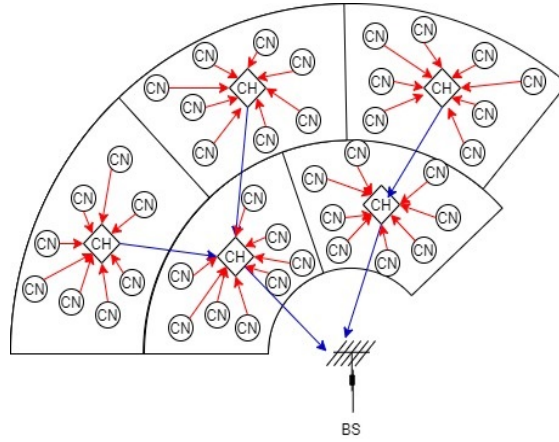


FIGURE 2.3: Gathering data from CNs and CHs

CHs. Fig. 2.3 shows an example of the process of collecting data: Nodes, other than CHs, are called Cluster Nodes (CNs).

CNs collect data on a regular basis and send it to CHs whenever requested. In turn, CH makes SS, if the channel is free, then it requests data from their CNs as well as from neighboring CHs. All CNs transmit their packets toward their appropriate CH. In order to avoid collision among CHs transmission, CSMA/CA and TDMA techniques can be used [85].

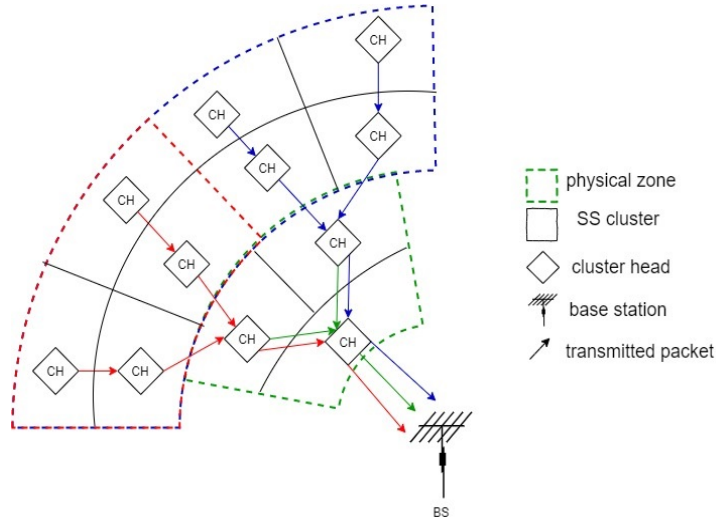


FIGURE 2.4: Data Routing between CHs located in different similar physical value zones

### 2.2.3 In network data aggregation

Generally, CNs and CHs close to each other may collect similar physical values, like humidity, temperature, etc. Fig. 2.4 shows a part of CRSN divided into clusters and virtually divided into physical zones with similar clusters (dashed lines) which have similar data packets. The network is virtually divided into physical rings, and contains  $k_p$  physical rings, each physical ring contains  $\frac{k}{k_p}$  physical zones where  $k_p \leq k$ , see Fig. 2.5. In average, each physical ring contains a number of physical zones, formed by an equal number of SS clusters. Let our CRSN has  $P$  physical zones generating in average  $P$  distinct packets in the entire CRSN. Each physical zone contains several SS clusters. Hence, the maximum number of distinct packets re-transmitted by the nearest CHs to the BS is  $\frac{P}{3}$ , where 3 is the number of clusters corresponding to the first ring (see Fig. 2.1). Fig. 2.4 shows that if a CH receives a  $p$  data packets from a neighbor CH located in the same physical zone, then it re-transmits  $p$  packets to the next CH. Whereas, if a CH receives  $p$  data packets from a CH in a different physical zone, then it re-transmits  $p + 1$  data packets to the forward CH [86], in other word, the similar data packet will be dropped.

*Proof.* See Appendix A.1. □

To detail the IDA process above, after collecting data coming from its CNs and neighboring CHs, a CH performs three steps to reduce the data redundancy, as described in our previous study [87]:

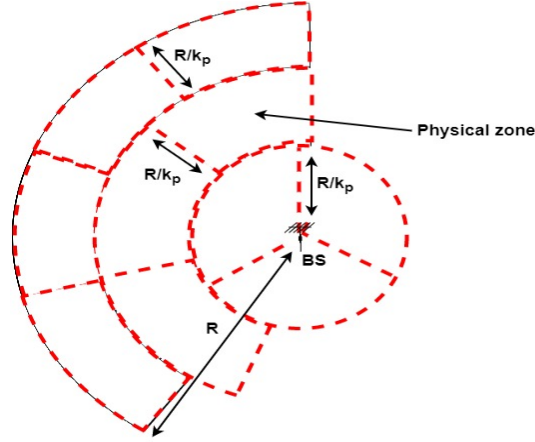


FIGURE 2.5: Physical zones in CRSN

1. CH aggregates the data received from CNs in a fixed size packet [87].
2. The CH applies a similarity test performing IDA process, based on the Euclidean's norm, applied to all packets.
3. Finally, a CH just transmits distinct packets towards the next CH (similar data packet will be dropped and not transmitted).

Obviously, CHs are prone to depletion of energy first; Therefore, the CH role is periodically rotated among potential nodes [59, 79]. The number of hops from the CH to the BS depends on the distance between them. The ring number in the CRSN can be considered as the maximum number of hops between the farthest CH and BS. A high number of hops means a low amount of energy will be consumed by the transmission, but the energy consumed by SS will be increased, and *vice versa*.

Hereinafter, we discuss how to find the optimal number of clusters that saves energy and improves the network lifespan, while satisfying an acceptable level of protection for the PU against the interference characterized by a high probability of PU signal detection, and achieving high spectrum efficiency by reducing the probability of false alarm on the PU presence.

## 2.3 Problem Formulation

In WSN, each node collects data from its environment and sends them toward CH. In our model, a CH has two major roles. The first one is performing the SS before requesting the data from the CNs and its previous CHs. The second one, by using IDA

concept, CH gathers and re-transmits data from its CNs and its backward CHs as per Ad-Hoc topology. Therefore, the main energy consumption of the CH comes from two operations. By performing the SS operation, an energy amount,  $e_s$ , is consumed. And when data is transmitted from a given CH towards the next one during data gathering. Notice that, the energy of data transmission depends on the decision outcome of SS performed by CH giving the permission to transmit the data: the CH may make a miss-detection decision on the PU status, i.e. PU is detected absent while it is truly active. Thus, a collision may take place and a re-transmission of data is required. Or the CH makes a true decision on the absence of PU. In this case, CH can send its data towards the next CH freely on the PU channel. Accordingly, the probability that a CH transmits the data is given by:

$$p_t = p(H_0)(1 - p_{fa}) + p(H_1)(1 - p_d) \quad (2.4)$$

where  $p(H_0)$  and  $p(H_1)$  are the prior probabilities that the channel is idle or active respectively.  $p_{fa}$  and  $p_d$  are the desired false alarm and detection probabilities respectively ensuring the SS requirements.

### 2.3.1 SS energy consumption and lifespan

The number of SS operations and the transmission range are directly related to the number of rings  $k$ . Therefore, finding the optimal number of rings  $k$ , which extends the network lifespan  $L$ , becomes essential for the network. Furthermore, network lifespan is directly affected by the energy consumption along CHs and CNs in our network. Because of the obvious difference in power consumption between the CHs and the CNs, our problem can be formulated as finding the optimal  $k$  according to energy consumed by the CHs only.

Let us define,  $e_i$ , the average energy consumed by a node located in the  $i_{th}$  ring:

$$e_i = \left[ e_r \left( \frac{R}{k} \right)^2 p_t N_{p_i} + e_s + e_0 N_{p_i} \right] \frac{M}{N} + \left[ e_r \left( \frac{R}{k} \right)^2 p_t + e_0 \right] \left( 1 - \frac{M}{N} \right) \quad (2.5)$$

where  $N_{p_i}$  is the average number of packets relayed by CHs in the  $i_{th}$  ring,  $e_r$  is the energy density at the receiver antenna and  $e_r \left( \frac{R}{k} \right)^2$  is the amount of energy consumed by CH during the data transmission.  $N$  is the total number of nodes and  $\frac{M}{N}$  denotes the probability of a node to come a CH, whereas,  $(1 - \frac{M}{N})$  is the probability of a node to be a CN.  $e_0$  is the energy consumed by the node in the running mode.

Let us define  $e_{max}$  as the maximum consumed energy per CHs in the CRSN network. In this work, the optimization is done with respect to the First Node Dies (FND) criterion (the simulation stops as soon as any node runs out of power). Lifespan is assumed to be



the number of iterations which can be done by the CRSN before the extinction of the first CH in the network [88].

Assuming each node is initially provided with an amount of energy called the initial energy,  $e_{int}$ . Hence, Lifespan  $L$  can be calculated as follows:

$$L = \frac{e_{int}}{e_{max}} \quad (2.6)$$

The optimisation problem, looking for extending the network lifespan, can be defined as follows:

$$\begin{aligned} \max_{k \neq 0} \quad & \frac{e_{int}}{e_{max}} \\ \text{s.t.} \quad & M \leq N \end{aligned} \quad (2.7)$$

Since  $e_{int}$  is constant, the maximization of (2.7) can be reformulated as a minimization problem as follows:

$$\begin{aligned} \min_{k \neq 0} \quad & e_{max} \\ \text{s.t.} \quad & M \leq N \end{aligned} \quad (2.8)$$

Using Eq. (2.5), we substitute  $e_{max}$  by its value in Eq. (2.8) to obtain:

$$\begin{aligned} \min_{k \neq 0} \quad & \left[ \max_i \left( e_r \left( \frac{R}{k} \right)^2 p_t N_{p_i} + e_s + e_0 N_{p_i} \right) \frac{M}{N} + \left( e_r \left( \frac{R}{k} \right)^2 p_t + e_0 \right) \left( 1 - \frac{M}{N} \right) \right] \\ \text{s.t.} \quad & M \leq N \end{aligned} \quad (2.9)$$

It becomes obvious that the closest CHs to the BS (ring 1) consumes the largest amount of energy due to the relaying role. Hence, our problem can be reformulated as :

$$\begin{aligned} \min_{k \neq 0} \quad & \left[ \left( e_r \left( \frac{R}{k} \right)^2 p_t k_p^2 + e_s + e_0 k_p^2 \right) \frac{M}{N} + \left( e_r \left( \frac{R}{k} \right)^2 p_t + e_0 \right) \left( 1 - \frac{M}{N} \right) \right] \\ \text{s.t.} \quad & M \leq N \end{aligned} \quad (2.10)$$

*Proof.* See appendix A.3. □

### 2.3.2 Optimal number of clusters

The optimal number of clusters allows us to extend the lifespan of our network, thus, by replacing  $M$ , i.e. the number of CHs in the network, by its value  $M = 3k^2$  as seen in section 2.2; then the optimization problem can be rewritten as follows:

$$\min_{k \neq 0} \left[ \left( e_r \left( \frac{R}{k} \right)^2 p_t k_p^2 + e_s + e_0 k_p^2 \right) \frac{3k^2}{N} + \left( e_r \left( \frac{R}{k} \right)^2 p_t + e_0 \right) \left( 1 - \frac{3k^2}{N} \right) \right] \quad (2.11)$$

s.t.  $M \leq N$

**Theorem 2.1.** *If  $f(x)$  is a uni-modal function defined over a real set  $\mathcal{G} \subset \mathbb{R}^+$  having a minimum at  $x_0$ , then,  $f(n)$ ,  $n \in \{\mathbb{N} \cap \mathcal{G}\}$ , is a set of points having a minimum  $M = \min\{f(\lfloor x_0 \rfloor), f(\lceil x_0 \rceil)\}$ .*

where  $\lfloor \cdot \rfloor$  and  $\lceil \cdot \rceil$  stand for the floor and ceiling operator respectively.

*Proof.* Let  $f(x)$  be a uni-modal function defined on  $\mathcal{G} \subset \mathbb{R}^+$  with  $\frac{df}{dx} \Big|_{x=x_0} = 0$  and  $\frac{d^2f}{dx^2} \Big|_{x=x_0} > 0$ . Let  $n \in \{\mathbb{N} \cap \mathcal{G}\}$  be a natural variable and  $U_n = \{f(n) \mid n \leq x_0\}$  and  $V_n = \{f(n) \mid n \geq x_0\}$  are two numerical sequences. knowing that  $U_n$  is monotonically descendant, and  $V_n$  is monotonically ascendant, thus,  $f(\lfloor x_0 \rfloor)$  and  $f(\lceil x_0 \rceil)$  are the minimum of  $U_n$  and  $V_n$  at  $\lfloor x_0 \rfloor$  and  $\lceil x_0 \rceil$  respectively. Obviously, the min of  $f(n)$  is  $M = \min\{\min\{U_n\}, \min\{V_n\}\}$ , hence,  $M = \min\{f(\lfloor x_0 \rfloor), f(\lceil x_0 \rceil)\}$ .  $\square$

Using the theorem 2.1, we can show that our Eq. (2.11) can be associated to an uni-modal function with a single minimum on its defined domain. Thus, our problem can be simplified into a simple optimisation problem. Replacing  $k$  by a continuous variable  $x$  in the Eq. (2.11), and differentiating  $e_i$  with respect to  $x$ , we obtain the optimal value:

$$k_{opt} = \operatorname{argmin} \left( e \left[ R^4 \sqrt[4]{\frac{e_r p_t \rho}{e_s + e_0(k_p^2 - 1)}} \right], e \left[ R^4 \sqrt[4]{\frac{e_r p_t \rho}{e_s + e_0(k_p^2 - 1)}} \right] \right) \quad (2.12)$$

if  $k_{opt} \gg 1$ , then we can approximate  $k_{opt}$  by:

$$k_{opt} = \left\lfloor R^4 \sqrt[4]{\frac{e_r p_t \rho}{e_s + e_0(k_p^2 - 1)}} \right\rfloor \quad (2.13)$$

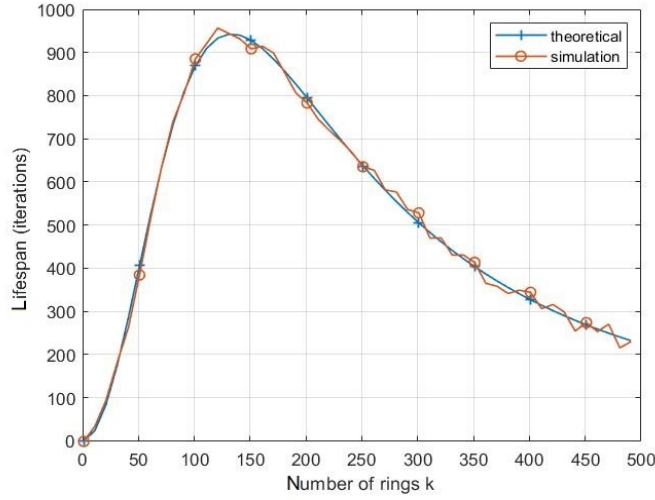


FIGURE 2.6: Theoretical and simulation results of the lifespan vs number of rings in CRSN

where  $\rho = \frac{N}{\pi R^2}$  is the density of the network. Since FND criterion is adopted in this work, the optimal value  $e_{opt}$  of  $e_{max}$  is given by substituting  $k_{opt}$  in Eq. (2.5):

$$e_{opt} = 2\sqrt{\frac{e_r p_t e_s}{\rho}} + \left(k_p^2 - 1\right) \left(e_0 \sqrt{\frac{e_r p_t}{e_s \rho}} + \frac{e_r p_t}{\rho}\right) \quad (2.14)$$

Figure 2.6 shows the lifetime of the nearest node with respect to the number of clusters. As it can be seen, the theoretical results coincide with the simulated ones approving Eq. (2.5). Lifespan increases until reaching a maximum before decreasing again. Indeed, when the number of clusters is low, the size of a cluster becomes large, then nodes consume a lot of transmission energy. Subsequently, the energy consumed while transmission becomes greater than the energy consumed to perform the SS by CH. As the number of clusters increases, their size decreases, hence, the transmission energy consumed by CHs decreases. Therefore, the lifespan increases until reaching a maximum value at a specific value of  $k$ . However, if the number of cluster continues decreasing, the energy consumed by SS is considerable high compared to the low transmission energy, so the lifespan decreases.

The number of clusters is related directly to the number of rings  $k$  as shown in Eq. (2.3). Hence, the optimal value  $k_{opt}$  of  $k$  should maximize the lifespan of the network.

## 2.4 Network Latency and Collision Probability

In this section, we derive the average time taken by each round of data gathering in the network, and the probability of collision with the PU. Furthermore, using the analytical

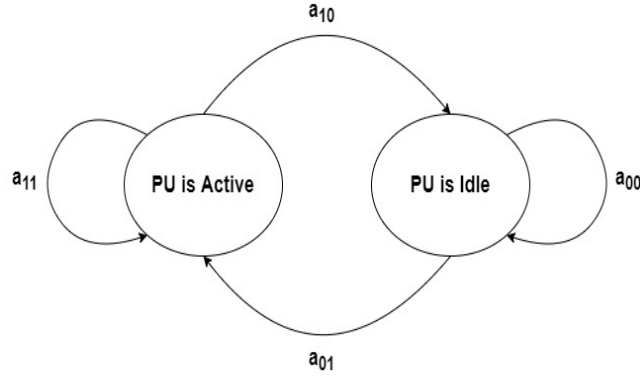


FIGURE 2.7: The PU activity modelled as a two-state Markov chain

expressions of the collision and the latency, we prove analytically that the first ring ( $k = 1$ ) of the network spends the larger amount of energy across the network.

#### 2.4.1 Collision Probability and Latency

Each time the BS requests data from nodes, the CNs start sending to the CHs towards the BS. In turn, before receiving data from CNs and backward CHs, CH should sense the channel in order to ensure that PU is absent to avoid any collision.

The state of PU (Active or Idle) is modelled by a discrete Markov chain as shown in Fig. 2.7 [89], where  $a_{ij}$  is the transition probability of the PU from the state  $i$  to the state  $j$ . The time in this Markov chain is divided into equal slots of  $\tau_s$  seconds.

On the other hand, each CH, acting as SU, has four tasks to do:

1. Performing SS to avoid any collision with PU.
2. Broadcasting a message to CNs and neighboring CHs indicating that it can receive collected data.
3. Gathering data from CNs and neighboring CHs and aggregating that data.
4. Waiting a broadcasting message from the next CH before transmitting packets toward it.

The data sending starts from the CHs of the farthest ring to the next one towards the BS. The probability that a collision occurs during the data transmission from the farthest ring till the BS is given by:

$$P_c = 1 - \prod_{i=1}^k (1 - P_{c_i}) \quad (2.15)$$

where  $P_{c_i}$  is the probability of collision with PU in the  $i$ th ring of a network. Such collision may take place in one of the following scenarios:

- (i) a miss-detection decision on the PU presence is made by the CH.
- (ii) the CH makes a correct decision on the PU absence, but the latter resumes its activity during the transmission period of the CH.

Consequently  $P_{c_i}$  could be evaluated as follows:

$$P_{c_i} = P(H_1) (1 - p_d^{m_i}) + P(H_0) (1 - p_{fa}^{m_i}) \left( 1 - a_{00}^{\frac{T_{a_i}}{\tau_s}} \right) \quad (2.16)$$

$a_{00}$  is the transition probability that the PU will stay in the idle state at the next time slot  $\tau_s$ , and  $T_{a_i}$  is the average time of transmission activity of CH in the  $i_{th}$  ring:

$$T_{a_i} = \tau_B + \tau_{G_i} \quad (2.17)$$

$\tau_B$  is the collecting data period from CNs,  $\tau_{G_i}$  is the period of data gathering from CNs members and previous CHs of a CH in the ring  $i$ :

$$\tau_{G_i} = \frac{NT_{CN}}{3k^2} + \tau_{T_i} V_i \quad (2.18)$$

where  $N$  is the total number of CNs in the overall CRSN,  $T_{CN}$  is the period of transmission data packet from a single CN to its appropriate CH, and  $\tau_{T_i}$  is the average time of collecting data from  $V_i$  CHs (in average) coming from the previous SS ring  $i + 1$ . Eq. (2.18) reflects the adoption of the Time Division Multiple Access (TDMA) technique by the CNs, as only one channel is assumed to be used by these nodes to send their data to CH. TDMA in our case prevents the interference between the transmissions of the CNs and lets the CH collect the data properly.

As mentioned previously,  $V_i$  is the average number of CHs which transmit data packets to a single CH in the ring  $i$ , in other words, the average number of previous CHs of any CH in the  $i_{th}$  ring in our CRSN. This value can be calculated as follows:

$$V_i = \frac{3(2(i+1) - 1)}{3(2i - 1)} = \frac{2i + 1}{2i - 1} \quad (2.19)$$

where  $i \in [1, k - 1]$ .

Neighbor CNs may collect similar data. Therefore, we assume that there are  $P$  distinct physical zones with similar areas. Thus, the average time of transmission  $\tau_{T_i}$  of each CH member of ring  $i$  can be presented as follows:

$$\tau_{T_i} = \frac{T_P}{m_i} \sum_{j=\lceil \frac{k_p}{k} i \rceil}^{k_p} \sum_{t=1}^{m_i} \left( \left\lceil \frac{mp_j}{m_i} t \right\rceil - \left\lfloor \frac{mp_j}{m_i} (t-1) \right\rfloor \right) \quad (2.20)$$

where  $\lceil \cdot \rceil$  stands for the ceiling operator,  $\lfloor \cdot \rfloor$  stands for the floor operator,  $T_P$  is the period of a single packet transmission,  $k_p$  is the number of physical ring,  $m_{p_j}$  is the number of physical zones in a physical ring  $j$  and  $m_i$  is the number of SS clusters in the  $i_{th}$  ring.

Eq. (2.20) can be rewritten as:

$$\tau_{T_i} = \frac{T_P}{m_i} \sum_{j=\lceil \frac{k_p}{k} i \rceil}^{k_p} m_{p_j} + m_i - (m_{p_j} \wedge m_i) \quad (2.21)$$

with:

$$m_{p_j} + m_i - (m_{p_j} \wedge m_i) = \sum_{t=1}^{m_i} \left( \left\lceil \frac{mp_j}{m_i} t \right\rceil - \left\lfloor \frac{mp_j}{m_i} (t-1) \right\rfloor \right) \quad (2.22)$$

where  $(m \wedge n)$  stands for the greatest common divisor operator of  $m$  and  $n$ .

*Proof.* See Appendix A.2. □

Consequently, the period of overall collecting data cycle from the network becomes:

$$\begin{aligned} T_c &= \frac{1}{P(H_0)(1 - p_{fa})} \sum_{i=1}^k T_{a_i} \\ &= \frac{1}{P(H_0)(1 - p_{fa})} \sum_{i=1}^k \left[ \frac{NT_{CN}}{3k^2} + \frac{T_P V_i}{m_i} \sum_{j=\lceil \frac{k_p}{k} i \rceil}^{k_p} m_{p_j} + m_i - (m_{p_j} \wedge m_i) \right] \end{aligned} \quad (2.23)$$

In Eq. (2.23), we assume that any transmission between two CHs, or CNs and CHs cannot succeed in presence of PU; Therefore, we divide the term  $\sum_{i=1}^k T_{a_i}$  (the total time of collecting data without interruption of PU) over the probability of transmission data without presence of PU.

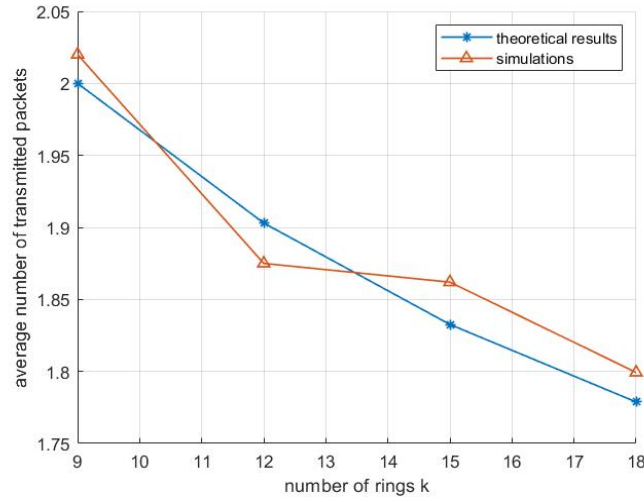


FIGURE 2.8: Comparison between the theoretical and the simulation result for the average number of relayed packets

### 2.4.2 Maximal Consumed Energy

Regarding Eq. (2.14) on  $e_{opt}$ , in order to prove that the CHs in the first ring consume the maximum amount of energy among all other CHs, we can instead prove that the CHs in that ring relay the maximum number of packets toward the BS. The average number of packets relaying by CHs in the  $i_{th}$  ring,  $N_{p_i}$  can be derived from Eq. (2.21):

$$N_{p_i} = \frac{1}{m_i} \sum_{j=\lceil \frac{k_p}{k} i \rceil}^{k_p} (m_{p_j} + m_i - (m_{p_j} \wedge m_i)) \quad (2.24)$$

Fig. 2.8 shows the comparison of the average number of relayed packet by a node in the middle ring between the simulation and the theoretical result. Each CH in the ring  $i$  relay  $N_{p_i}$  packets, in this case, a CH at the first ring should aggregate  $N_{p_1}$  packets:

$$N_{p_1} = \frac{1}{3} \left( \sum_{j=1}^{k_p} 3(2j-1) + 3 \sum_{j=1}^{k_p} 1 - \sum_{j=1}^{k_p} (3(2j-1) \wedge 3) \right) = k_p^2 \quad (2.25)$$

Hereinafter, we show that:

$$N_{p_1} > N_{p_i}, \quad \forall i > 1. \quad (2.26)$$

*Proof.* See Appendix A.3.

□

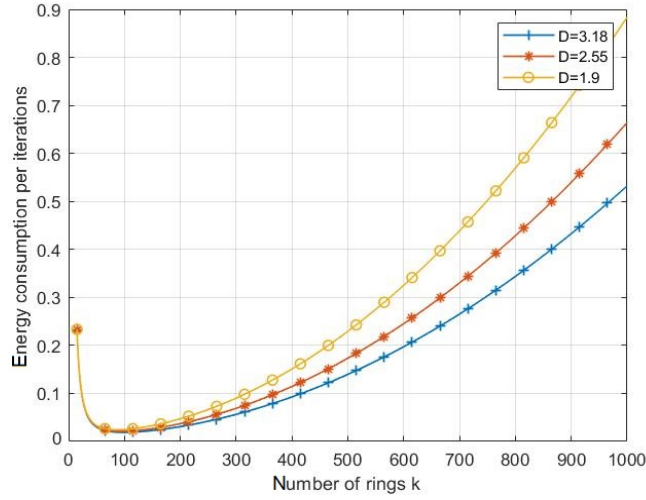


FIGURE 2.9: The effects of the ring number on the energy consumption per iteration on  $e_{max}$  with different network densities

## 2.5 Evaluation of our system

In this section, we investigate our theoretical results and show the effectiveness of the IDA in reducing both the collision probability and the network latency. The results are compared with the classical CR without IDA techniques. In our simulations, we consider a circular network containing  $N$  nodes with a radius  $R = 100\text{ m}$  and 3 physical rings. Thus, the number of physical zones is  $P = 27$ . The processing energy of nodes during the receiving is  $e_0 = 50\text{ nJ}/\text{packet}$  and the transmission energy constant is  $e_r = 0.1\text{ nJ}/\text{bit}/\text{m}^2$  [90]. For the sensing process, a SNR of  $-1\text{ dB}$  is considered, and  $p_d = 0.9$  and  $p_{fa} = 0.1$  are set as the target probabilities to meet the SS requirement. An Energy Detection method [49] is considered for SS, and the consumed energy by the sensing  $e_s$  is evaluated according to the number of samples  $N_s$  needed for the sensing process [91] and the energy consumed per sample  $E_s$ :

$$e_s = N_s E_s \quad (2.27)$$

Fig. 2.9 depicts the effects of the ring number  $k$  on the energy consumption by iteration with different values of node density  $D$ . It can be shown that increasing the density of nodes inside the network leads to decrease the energy consumed during an iteration. As for the impact of the number of rings  $k$ , the energy consumption per iteration decreases with the increase of  $k$  until the latter reaches an optimal value, at which this energy consumption becomes minimal. If  $k$  continues increasing, the energy consumption per



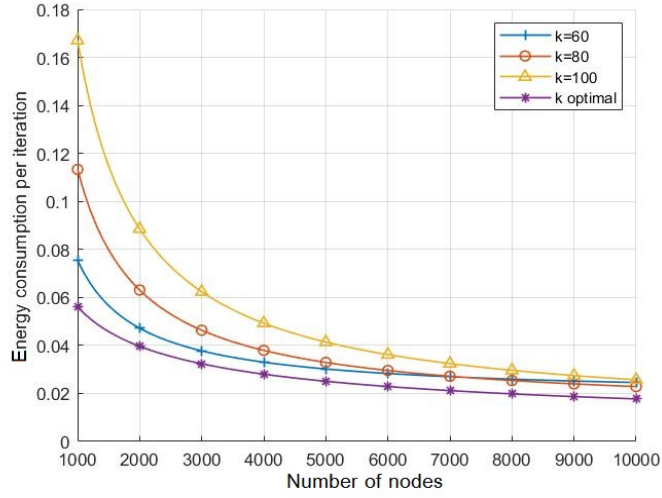


FIGURE 2.10: The effects of the number of nodes on the energy consumption with different values of  $k$

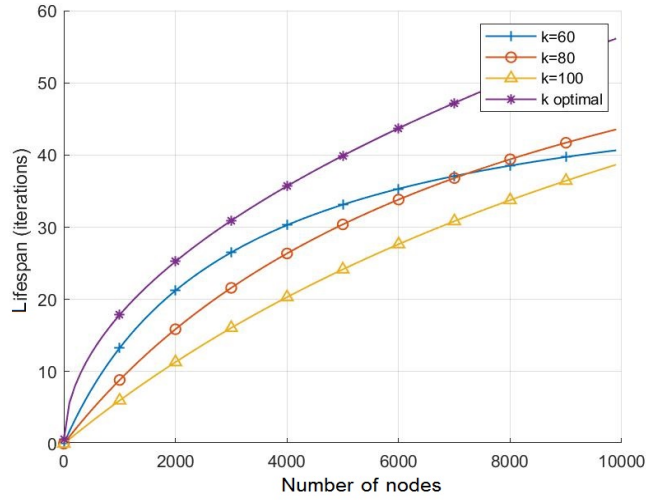


FIGURE 2.11: The effects of the number of nodes on Lifespan with respect to various values of  $k$

iteration re-increases leading to shortening the network lifespan.

Figures 2.10 and 2.11 highlight the effects of the nodes number on the energy consumption and lifespan respectively for various values of  $k$  including the optimal value. As the two figures show, for all the considered values of  $k$ , the energy consumption per iteration decreases with respect to the number of nodes while the network lifespan increases. However, when setting  $k$  to its optimal value for each node number  $N$ , Fig. 2.10 shows the energy consumption per iteration decreases compared to the other considered values of  $k$ , which leads to extend the network lifespan as presented by Fig. 2.11.

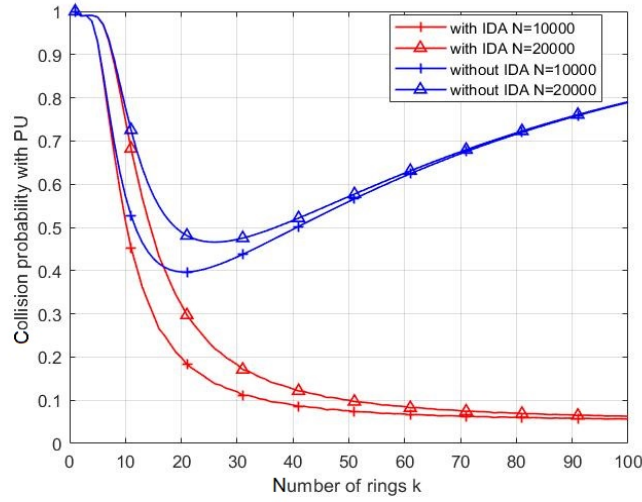


FIGURE 2.12: The collision probability of CH with PU during its activity with and without IDA.

Hereinafter, we investigate the performance of the CRSN with and without IDA in terms of the collision probability and the data delivery latency. The numerical results are based on the data size effect. Indeed, other message contents, such as the header and the trailer of the message, may affect the performance, but these contents are not considered in our work.

The collision probability of CHs with PU in the middle ring (ring  $\lceil \frac{k}{2} \rceil$ ) of the network during the PU activity, with and without IDA are shown in Fig. 2.12. For both cases (with and without IDA), the collision rate increases with  $N$  due to the increasing number of messages to be transmitted. When IDA is applied, the probability of collision with PU decreases with respect to the number of rings. This effect is due to both the aggregation of the messages and the reduction of the number of CNs in each cluster. Hence, the time of collecting data in each cluster decreases leading to decrease the probability of collision. By contrast, without IDA, the probability of collision increases after reaching a minimum value. This refers to the huge number of packets re-transmitted by each CH coming from the previous CHs. Without IDA, the packet length increases linearly with the number of rings and results in long transmission time of CHs.

Fig. 2.13 shows the duration of complete collecting data cycle from CN to the BS. In comparison to the case without IDA, IDA allows the CRSN to significantly reduce the collecting data cycle period. The CH spends a significant amount of time collecting data from its cluster nodes in all cases and for a small number of rings. This time decreases

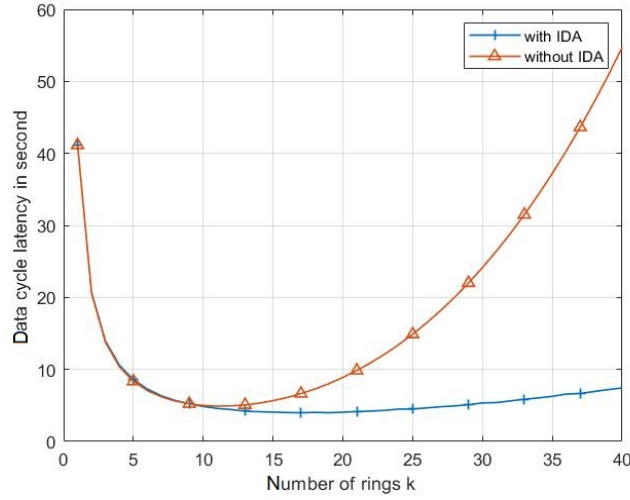


FIGURE 2.13: The duration of complete collecting data cycle from CN to the BS with and without IDA

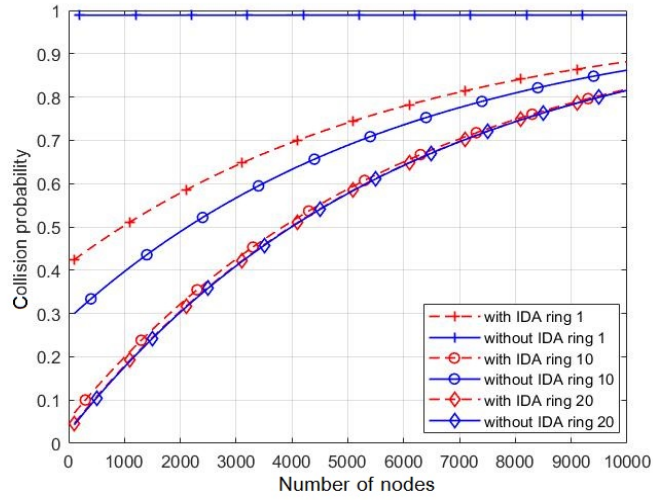


FIGURE 2.14: The collision probability of CH with PU during its activity with and without IDA with different rings

as  $k$  increases until it reaches a predetermined value. Collecting time increases once the value of  $k$  is exceeded. This effect is caused by multi-hop transmission, in which a large number of routes are created as a result of high  $k$  values, resulting in a long transmission time between CHs and BS.

Fig. 2.14 depicts the collision probability with and without IDA during a transition of rings:  $k = 1, 10$  or  $20$ , with respect to the number of nodes in the network. Clear effect of the IDA can be noticed on reducing the collision compared to case of no IDA for the three considered values of  $k$ . For instance, for  $k = 1$  and  $N = 1000$ , the collision rate is about 0.5 when IDA is adopted, whereas it is closed to 1 for no IDA. However, the

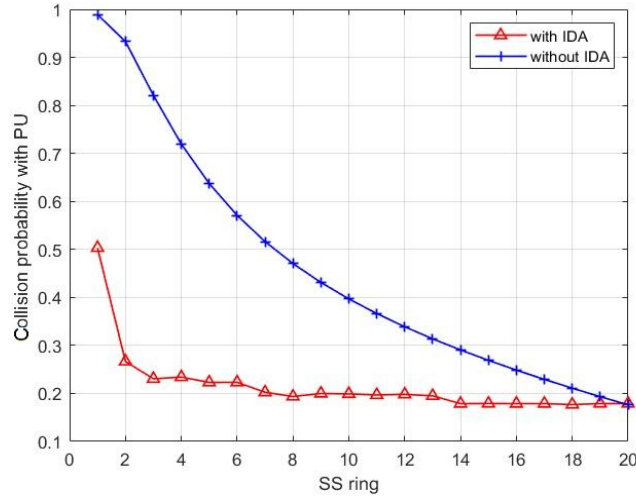


FIGURE 2.15: The collision probability of CH with PU during its activity among all network rings

collision rate exhibits an increasing with  $N$  for both cases due to the increase in the number of messages to send.

Fig. 2.15 shows the collision probability in networks, with and without IDA, in different rings where the number of nodes is fixed to  $N = 1000$ . It can be shown that, at the nearest ring from the BS (ring 1), we could notice a big difference between the two curves due to the redundancy reduction of the message obtained when IDA is applied. This must reduce the message length and, thus, alleviate the collision rate. Moving forward to the farthest ring from the BS (in our example it is ring 20), the curves begin to converge until they intersect at the last ring. This is because the collision with PU at the last ring during the SU activity is the same for both cases, with and without IDA since there are no packets coming to the last ring, i.e. the CHs at this ring do not relay any packet from previous rings.

Fig. 2.16 shows a comparison between our system and the Distributed Spectrum-Aware Clustering (DSAC) system described in [69]. By taking into consideration both the maximum transmission distance between sensor nodes as well as the density of sensor nodes in the network, the authors of DSAC were able to determine the optimal number of clusters to be used in a network. It is important to note that our simulations revealed that the energy usage of our system is much lower than that of DSAC system. The reason behind this is that when we optimise our problem, we take into account the amount of energy consumed during the SS process.

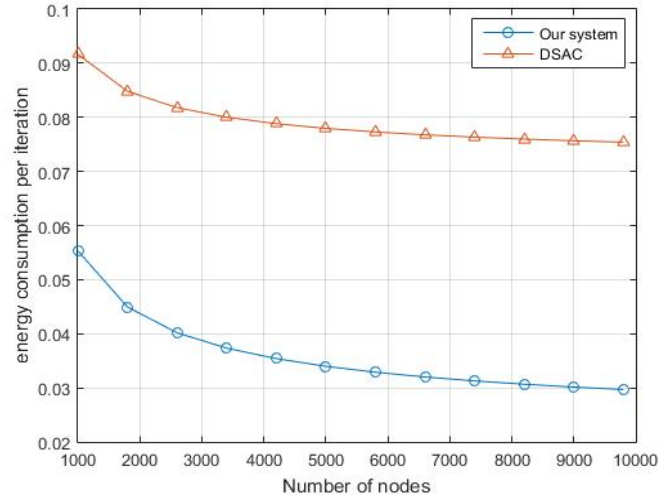


FIGURE 2.16: Comparison of energy consumption between our system and the DSAC

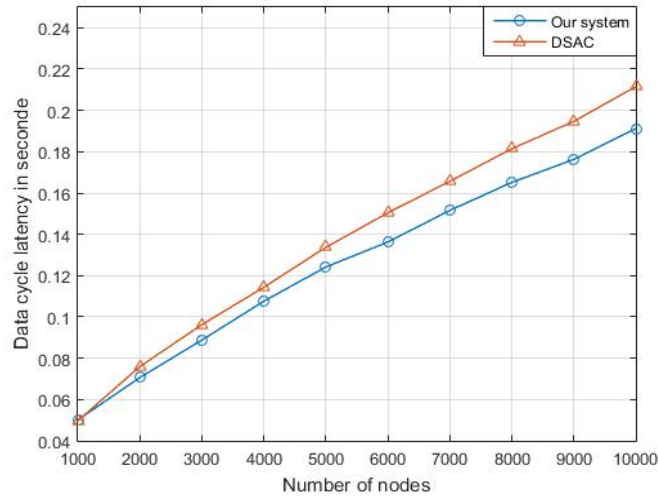


FIGURE 2.17: Comparison of data latency between our system and the DSAC

Additionally, we compare the latency (Fig. 2.17) and collision probability with the PU (Fig. 2.18) of our system's entire network to the DSAC. Our system beats the DSAC in terms of collision probability and delay, as demonstrated by the simulation.

## 2.6 Conclusion

In this study, we consider a CRSN in which all nodes have access to the SS module. This network is divided into clusters, each with a number of nodes and a single CH. A licensed channel dedicated to PU is used for communication between CHs and the base station. The CHs are in charge of SS. When CH determines that PU is not present, it

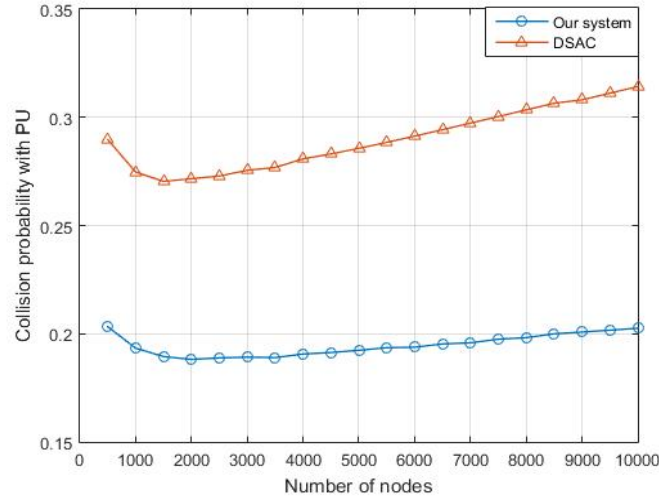


FIGURE 2.18: Comparison of collision with the PU transmission between our system and the DSAC

sends a request to its cluster's nodes and backward CHs to gather data.

A low number of clusters in this network increases energy consumption due to data transmission, whereas a high number of clusters increases energy consumption due to SS. As a result, we calculate the optimal number of clusters, which reduces CRSN's energy consumption and increases the network's lifespan. After calculating the average number of packets re-transmitted by each CH, latency in delivering the message and collision rate between the CH and PU transmission are also calculated given the IDA adoption by the CRSN. The use of IDA and the selection of an optimal cluster number can significantly extend the network lifespan, according to numerical results.

In the next chapter, we will focus on multi-hop routing protocols in WSNs, we propose a new routing protocol called LIBRO that needs a little processing capability compared to other protocols, and can also be applied to mobile networks, as the nodes in this protocol do not need to know the identities of neighboring nodes neither their geo-locations. Thus, the memory space that this protocol allocates will be small as well.

## Chapter 3

# A new routing algorithm for WSN (LIBRO)

The nodes in Wireless Sensor Networks (WSN) can be distributed at random across a harsh geographic zone. These nodes are usually powered by batteries and have limited transmission and processing capabilities. Because replacing a battery may not always be possible, managing the residual energy in such a network is critical. Energy is essentially consumed during packet transmission phase. As a result, routing protocols become extremely important because they have an impact on data transmission energy consumption. In this chapter, we introduce a new routing protocol for uplink multi-hop WSNs that is based on geographical location information. The proposed protocol assumes that the geographical zone and the data transmitted are more valuable than the source node ID. Without knowing the topology or path nodes between the source and the destination, our proposed protocol ensures data packet delivery in a dense network. The average consumed energy, packet loss probability, and mean delivery time are all calculated analytically. In terms of connectivity, lifespan, memory, and latency, numerical results show that our protocol outperforms the industry-standard Distance Routing protocol (DIR).

### 3.1 Introduction and related works

Wireless Sensor Networks (WSNs) are made up of a collection of small devices known as nodes that are typically spread out over a large geographical area [6]. WSN nodes [7] are battery-powered devices with low-energy Radio Frequency (RF) modules and processing capabilities. Each node collects local physical data, which it then sends to a BS via a single or multi-hop route [8].

The efficient coverage of the geographical zone under surveillance, as well as full access to all nodes, so that data stored in any node can be accessed at any time [92], are two key challenges in WSN. The first problem can be solved by deploying an arbitrary and uniformly large number of WSN nodes throughout the network's geographical zone. The second problem can be solved by using a routing protocol [93] that ensures reliable traffic between the source and the destination.

### 3.1.1 Some multi-hop routing protocol in literature

Because of the limited energy supply and available bandwidth in WSNs, routing is extremely difficult [94]. On-demand protocols for Ad-Hoc and multi-hop WSNs have gotten a lot of attention in the last two decades to ensure access to WSN nodes, and several protocols have been proposed. The Ad-Hoc On-Demand Distance Vector (AODV) routing protocol was originally designed for Ad-Hoc networks and has since been extended to WSN [95–97]. Setting up a transmission route at the start of a communication session is the main concept of AODV. Based on a routing table with an entry per destination, the route is maintained until the end of the session. These tables are kept in WSN nodes that route traffic from the source to the destination.

The authors of [98] proposed an Enhanced AODV (EAODV), which shares similar features to AODV. However, the new protocol differs primarily in the following ways:

- There is no need to exchange route information on a regular basis in order to achieve routes for all other host nodes.
- Instead of being aware of network topology structures, nodes only need to maintain routes among themselves and other nodes in their routing tables.

Dynamic Source Routing (DSR) is proposed as an on-demand routing protocol [99, 100], similar to AODV. It is based on the source routing protocol, which states that a packet sent from a source  $S$  to a destination  $D$  should contain all of the route's intermediate node information [101]. When a source node wants to send a packet to a destination, it looks in its routing cache for a route to the destination and uses an existing one. Otherwise, the source node starts the route discovery phase, which includes route request and route reply (RREP) messages. To avoid using a large routing table and to reduce the control header in each transmitted packet, the Zone Routing Protocol (ZRP) was proposed [102, 103]. Also, this protocol reduces the latency caused by a route discovery process [104]. ZRP creates a routing zone around each node, allowing all nodes within a  $k$  hop



distance of a node  $X$  to join it. In ZRP, two types of routing protocols are used: intra-zone routing for nodes within the routing zone, and inter-zone routing for nodes outside the routing zone. As a result, before sending a packet from a source  $S$  to a destination  $D$ ,  $S$  determines whether  $D$  is in its routing zone. If this is the case, the packet is sent to  $D$  right away using the intra-zone proactive routing protocol. Otherwise, the route is searched by the inter-routing reactive protocol.  $S$  sends a route request to its routing zone's border neighbours, who, in turn, look for  $D$  in their own routing zones. The process continues until it finds  $D$ , at which point it sends a route reply to  $S$  [103].

### 3.1.2 On demand geo-location routing algorithms

Some proposed routing protocols in the literature rely on node location information. In this context, location Aided Routing (LAR) protocol is an on-demand routing protocol based on the location information [105]. The nodes in LAR should be equipped with GPS modules or be aware of their location. In terms of route discovery and response, LAR is similar to DSR. The main difference is in the route discovery phase: in DSR, RREQ requests flood the network, whereas in LAR, a source  $S$  predicts the location information of its corresponding destination  $D$ . As a result,  $S$  floods a request zone defined by a rectangle containing  $S$  and  $D$  when it uses LAR. If  $D$  is not within the request zone, a node discards a receiving RREQ message [93].

### 3.1.3 TBF and DIR algorithm

Another location-based protocol proposed in [106] is Trajectory-Based Forwarding (TBF). This protocol requires a sufficiently dense network as well as each node's geographical location information [107]. In the packet, the source's specified trajectory is introduced. Each node receives the packet and re-transmits it to the next hop, which is its nearest neighbour toward the destination. Because the identities of the nodes along the trajectory are not included in the packet, nodes' mobility has no effect on the TBF trajectory. The authors of [108] propose a Distance Routing protocol (DIR) based on geographical location data. In DIR Fig. 3.1, an intermediate node  $I$  re-transmits the packet to one of its neighbours,  $X$ , satisfying the lowest angle XSD, along the trajectory connecting  $S$  to  $D$ .

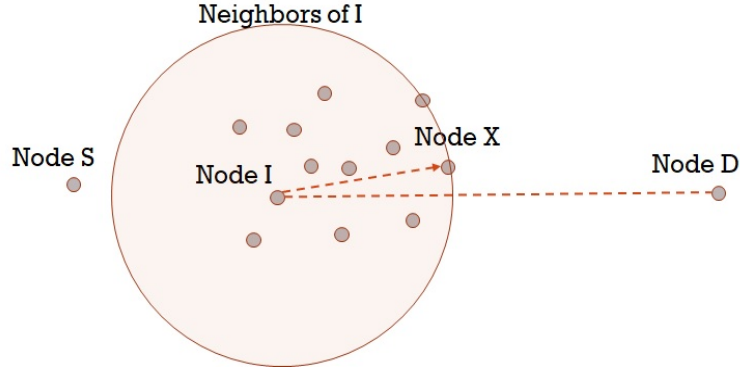


FIGURE 3.1: Distance Routing protocol an example

### 3.1.4 Presentation and contribution of LIBRO

Despite the fact that the aforementioned protocols rely on node location information, they do not emphasise the significance of the event location. In various scenarios, a user may want to collect data from a geographical zone of an event regardless of the identity of the source node. Motivated by this feature, we propose the Location Information-Based Routing protocol in this chapter (LIBRO). LIBRO focuses on the event location rather than the identity of the node. LIBRO routing works by drawing an imaginary line between the source  $S$  and the destination  $D$ . This line is a path that contains a collection of similar virtual circular zones called LIBRO zones that are evenly spaced along the path. By designating one node in each zone as a relay, all nodes within these zones participate in the routing process. Our protocol is multi-hop in nature, and a node within our WSN does not require any information about its neighbours. The packet is routed to a neighbouring geographical zone, where an arbitrary selected node within that zone re-routes it to the next zone. The other nodes in this zone return idle. When compared to other protocols such as LAR, TBF, and DIR, LIBRO requires less processing and memory to complete a routing process.

## 3.2 Location Information Based Routing Protocol (LIBRO).

The adopted WSN model consists of a large number of sensor nodes  $N$  [109], randomly distributed over a rectangular geographical area  $A$  with a node density  $\rho = \frac{N}{A}$ . Because the node density is assumed to be high, we can assume that any circular zone of area  $z = \pi R^2$  in this network contains at least one node, where  $R > 0$  is a pre-selected radius of LIBRO zones. As a result, a zone of area  $z$  contains  $\rho z$  nodes on average. The antenna gain at transmitter and receiver is 1. To save nodes wasted energy during the listening phase, a radio wake-up scheme is used to ensure that all nodes in regions of interest are

Notations List	
Notation	Definition
$N$	Total node numbers in the network
$A$	Area of geographical zone of the network
$\rho$	Network nodes density
$R$	Radius of LIBRO zone
$z$	Area of LIBRO zone
$w$	Maximum chosen delay
$t$	The current chosen delay
$n$	Number of steps between the Target and the Destination zones
$p$	Maximum fixed length of LIBRO step
$\tilde{p}$	Calculated step length where $\tilde{p} < p$
$\tilde{r}_t$	Exact transmission range of network node
$r_t$	Approximated transmission range of network node
$\tilde{e}_T$	Exact consumed transmission energy
$e_T$	Approximated consumed transmission energy
$e_r$	Energy density at the receiver antenna
$e_p$	Consumed energy during node processing
$E$	Total consumed energy of whole node cycle
$P_l$	Probability of data loss

TABLE 3.1: The notations.

listening and ready to receive and forward messages [110]. It is assumed that the nodes have a GPS module or are aware of their location. They are powered by a battery and include an initial energy  $e_{int}$ , a low-cost processor, a small memory capacity, physical sensing units, and a low-power transceiver. Furthermore, the WSN is assumed to be homogeneous, with all nodes capable of acting as routers. Nodes in our work may be mobile, allowing them to move within the network area. The WSN is linked to the internet via a gateway, which also serves as a base station (BS). Thus, the user requests specific data about a geographical zone from the WSN via the gateway, which gathers the data from the target zone and sends it back to the user via the internet.

The users can ask the network at any time for physical data from any geographical zone through internet connection that links the user to the gateway (see Fig. 3.2). The gateway plays the role of the base station in the WSN. In addition, to deal with the critical data event cases, another gateway can play the role of Static Base Station (SBS), which is installed at a location known by all nodes members. In this case, a node getting a critical data event, can transmit this data to the SBS using multi-hop transmissions.

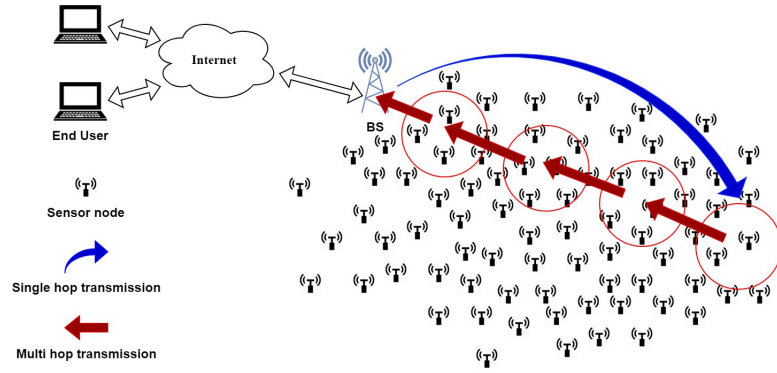


FIGURE 3.2: LIBRO uplink and downlink data transmission.

### 3.2.1 LIBRO contributions

We propose the LIBRO routing protocol to efficiently deliver data from a given geographical zone to BS. The proposed protocol can route data from the target zone to the destination using a multi-hop topology without knowing the identities of the nodes in between. This has a positive impact on the nodes' energy and memory efficiency. The main contributions of this protocol are as follows:

- Our new routing protocol collects and forwards the necessary data from a target geographical zone to the destination without the need for cooperation among the nodes on the routing path. In comparison to other previously published protocols, the node in our system does not require any prior knowledge of its neighbours' identities; it requires less memory in storing and processing the identities and locations of its neighbours.
- In our protocol, the packet header does not explicitly contain all of the hops between the source and the destination. As a result, the header has a small packet size.
- Unlike TBF and DIR, the proposed LIBRO ensures that the routing load is distributed fairly across a large number of nodes while transmitting data between a source and a destination. Each node accomplishes this by setting a random count-down delay. If a source wishes to send multiple packets to a destination, each packet may take a different trajectory, as opposed to TBF and DIR, which use the same nodes along the trajectory.
- Because nodes in LIBRO lack knowledge of the network topology, node mobility has no effect on the routing process. In order to transmit a data packet, a sensor

node sends its message to a LIBRO zone via a routing trajectory, where the presence of one node within the LIBRO zone is sufficient to ensure that the message is resent and the routing process is completed.

- LIBRO is simple to integrate into any mobile or static WSN. It does not require any special network infrastructure; all nodes in such a network are anonymous, and LIBRO does not require node IDs.

### 3.2.2 Presentation of our protocol

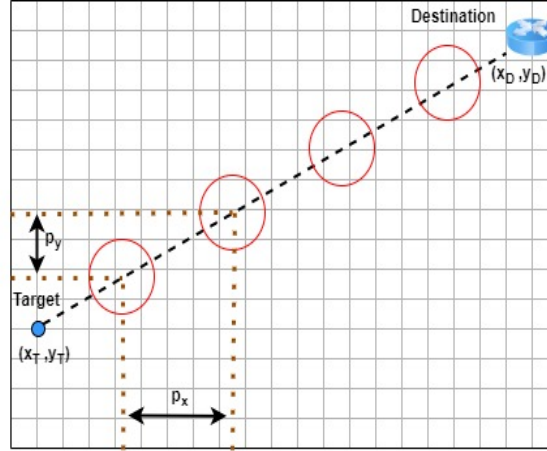
As shown in Fig. 3.2, LIBRO first divides the routing traffic into separate circular hop zones with equal radius  $R$ . As a result, hops are made between zones rather than nodes. As a result, the data packet travels through the intermediate zones from the source (target zone) to the destination (gateway). The current zone is the one that contains a node holding the packet before re-transmitting it to a subsequent zone toward the destination (BS). To illustrate our protocol, consider a user who wishes to obtain physical data from a specific target geographical zone  $T$  of the area  $\pi R^2$ . To that end, it uses the internet connection to request data from the gateway. Over the entire network, the gateway broadcasts a wake-up message and a data request packet containing the target zone coordinates  $x_T$  and  $y_T$ . Only the nodes within the target zone construct a physical data packet and then initiate a random countdown delay between 0 and  $w$ , where  $w$  is the maximum available delay; all other nodes outside the target zone go to sleep. The first node to exceed its waiting period (the countdown reaches zero) becomes a target node, at which point it draws a virtual line between its zone location  $T$  and the gateway location  $D$ . The coordinates for this line are  $(x_D, y_D)$  and  $(x_T, y_T)$ . The line is then divided into  $n$  equal steps (equal segments), and the distance between two consecutive steps are equal. The centre of each zone is located at the end of each segment. To summarise, a number of zones that are evenly distributed and spaced along the line that connects the source and the gateway is determined by the source node. Following that, the number of steps  $n$  can be expressed as follows:

$$n = \left\lceil \frac{\sqrt{(x_T - x_D)^2 + (y_T - y_D)^2}}{p} \right\rceil \quad (3.1)$$

where  $\lceil \cdot \rceil$  stands for the ceiling function and  $p$  is the distance between two consecutive steps, which is also the transmission range of a node in the network. After calculating the number of steps  $n$ , we can find the length of steps in  $x$  and  $y$  directions as  $p_x$  and  $p_y$  respectively (see Fig. 3.3).  $p_x$  and  $p_y$  are calculated as follow:

$$p_x = \frac{x_T - x_D}{n}, \quad (3.2)$$

$$p_y = \frac{y_T - y_D}{n} \quad (3.3)$$

FIGURE 3.3:  $x$  and  $y$  steps.

### 3.2.3 LIBRO Packet header and flowchart

The target node within a target zone  $T$  constructs its packet as shown in Fig. 3.4, where  $x_i$  and  $y_i$  represent the  $x$  and  $y$  locations of the centre of the next routing zone in the direction of the gateway, respectively. The control field indicates whether the packet is a data request packet (one-hop downlink transmission) or a response packet containing physical data sensing (multi hop uplink transmission). The target node then sends a data packet to the next zone. After receiving the target node transmission, the neighbours of this target node within the target zone drop their packets. In this work, the energy consumed by the other nodes within the target zone is regarded as negligible. The process then starts over at the next zone and continues until the data packet reaches the gateway. Our protocol's flowchart is shown in Fig. 3.5. The message will be received by all nodes within the target node's transmission range after the packet is broadcast by the target node. The first two fields ( $x_i$  and  $y_i$ ), which specify the current routing zone, are read by a node  $M_i$  with coordinates  $(x_{M_i}, y_{M_i})$  that receives the message (next hope). Then,  $M_i$  computes its  $d$  distance from the current zone centre  $(x_i, y_i)$ :

$$d = \sqrt{(x_{M_i} - x_i)^2 + (y_{M_i} - y_i)^2} \quad (3.4)$$

if  $M_i$  is outside the current zone, *i.e.*  $d > R$ , then, the node drops the packet. Other-

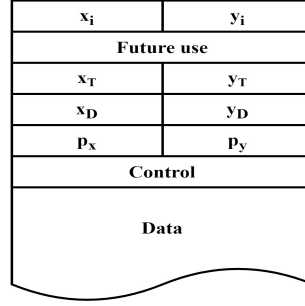


FIGURE 3.4: LIBRO packet header.

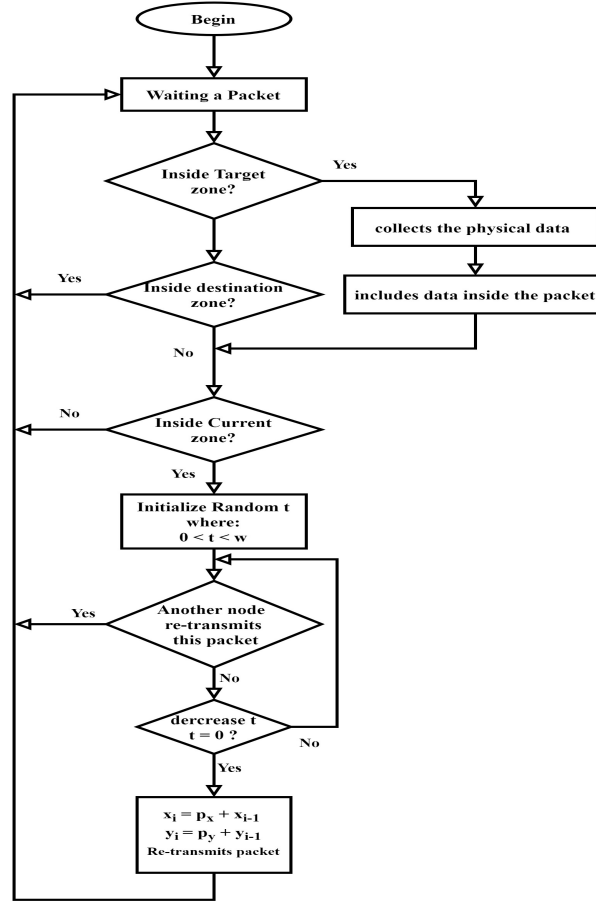


FIGURE 3.5: LIBRO flow chart.

wise, *i.e.*  $M_i$  is inside the current zone,  $M_i$  checks if the current zone is the destination zone. If not,  $M_i$  initialises a local random countdown delay  $t_i$  between 0 and  $w$ . If another node re-transmits the packet before the expiration of  $M_i$ 's delay, then  $M_i$  drops the packet. If the local delay is expired, and  $M_i$  does not receive any re-transmission in its current zone,  $M_i$  updates the current zone coordinates  $x_i$  and  $y_i$ , and re-transmits the packet. This cycle continues until reaching the destination zone where the gateway

exists where, all nodes inside destination zone will drop the message excepts the gateway. After that, the gateway re-transmits data packet to the user via internet.

### 3.2.4 LIBRO energy consumption

The amount of energy consumed by a network is critical in determining its lifespan. LIBRO, as illustrated in Fig. 3.2, is based on zones rather than nodes. As a result, instead of nodes, the packet is sent along the circular zones. Each packet may take a different path from  $S$  to  $D$  via various nodes. As a result, the load is distributed equally among all nodes within the zones along the traffic between  $S$  and  $D$ .

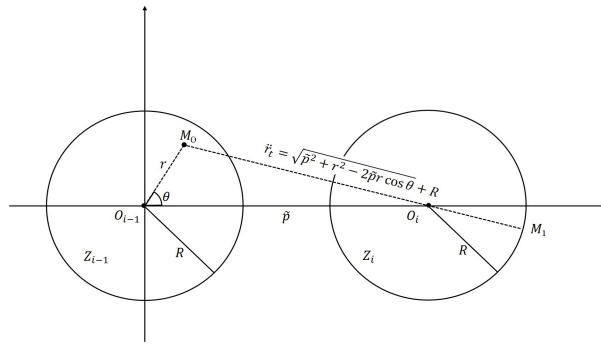


FIGURE 3.6: The exact transmission range of  $M$  which cover the whole next zone  $Z_i$ .

Ideally, a node located at point  $M_0$  in a previous LIBRO zone  $Z_{i-1}$  must cover the whole next LIBRO zone  $Z_i$  with its transmission range. Fig. 3.6 shows the distance between a node located at  $M_0$  in a previous zone  $Z_{i-1}$ , and the center of the current zone  $Z_i$  plus the LIBRO zone radius  $R$ . We assume that the center of  $Z_{i-1}$  is located at the origin of an orthonormal basis. The center of  $Z_i$  is located at  $x$ -axis and has as  $x$  coordinate  $\tilde{p}$ , where  $\tilde{p}$  is the distance between  $O_{i-1}$  and  $O_i$ , it is less or equals to the max step  $p$  of a node.

$$\tilde{p} = \frac{\sqrt{(x_T - x_D)^2 + (y_T - y_D)^2}}{n} = \sqrt{p_x^2 + p_y^2} \leq p \quad (3.5)$$

The transmission range  $\ddot{r}_t$  of node  $M$  is given by:

$$\ddot{r}_t = \sqrt{\tilde{p}^2 + r^2 - 2\tilde{p}r \cos \theta} + R \quad (3.6)$$



where,  $\theta$  is the angle between  $M$  and  $x$  axis, and  $r$  is the distance of  $M$  from the center of basis. Assuming a free-space transmission, the average consumed energy  $\ddot{e}_T$  related to  $\ddot{r}_t$ , to transmits a packet between two consecutive zones, is calculated as follow:

$$\begin{aligned}\ddot{e}_T &= \frac{e_r}{\pi R^2} \int_0^R \int_0^{2\pi} \ddot{r}_t^2 r d\theta dr \\ &= \frac{e_r}{\pi R^2} \int_0^R \int_0^{2\pi} \left( \sqrt{\tilde{p}^2 + r^2 - 2\tilde{p}r \cos \theta} + R \right)^2 r d\theta dr\end{aligned}\quad (3.7)$$

where  $e_r$  is the energy density needed by the receiver to properly decode a received message. The value  $r$  appeared in the double integral because we moved to the polar coordinate. We multiply the integral by  $\frac{1}{\pi R^2}$  because of that  $M$  is inside the LIBRO zone of area  $\pi R^2$ .

*Proof.* A simulation in java is made in order to proof this double integral, see appendix B.1.  $\square$

As we can see, the integral in Eq. (3.7) does not has a trivial solution, thus, we make a modification on transmission range  $\ddot{r}_t$  as shown in Fig. 3.7, where the node at  $M_0$  transmits its packet with a transmission range  $r_t$  reaching the outer intersection of the next zone with the  $x$  axis  $M_1$ . Let  $r_t = M_0M_1$  is an approximation value of  $\ddot{r}_t$ , where  $r_t$

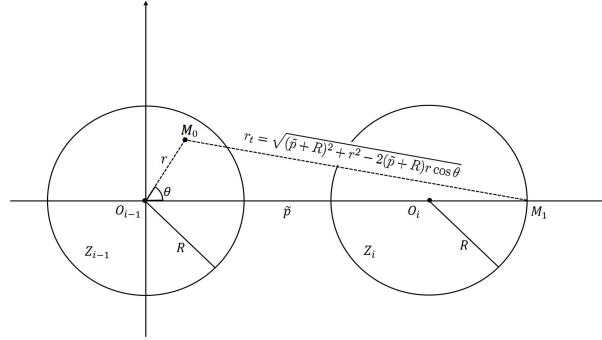


FIGURE 3.7: The distance between a node  $M$  in a previous zone,  $Z_{i-1}$ , outer intersection of the next zone  $Z_i$  with the  $x$  axis.

is calculated as follow:

$$r_t = \sqrt{(\tilde{p} + R)^2 + r^2 - 2(\tilde{p} + R)r \cos \theta} \quad (3.8)$$

Obviously,  $r_t \leq \ddot{r}_t$ . And the average consumed energy  $e_T$  related to  $r_t$ , to transmits a packet between two consecutive zones, becomes:

$$\begin{aligned} e_T &= \frac{e_r}{\pi R^2} \int_0^R \int_0^{2\pi} \ddot{r}_t^2 r d\theta dr \\ &= \frac{e_r}{\pi R^2} \int_0^R \int_0^{2\pi} ((\tilde{p} + R)^2 + r^2 - 2(\tilde{p} + R)r \cos \theta) r d\theta dr \end{aligned} \quad (3.9)$$

*Proof.* A simulation in java is made in order to proof the approximation between equations Eq. (3.7) and Eq. (3.9), see appendix B.1.  $\square$

Hence, we can demonstrate that:

$$\frac{e_T}{\ddot{e}_T} \geq 94\% \quad (3.10)$$

*Proof.* See appendix A.4.  $\square$

Therefor,  $e_T$  is a good approximation of  $\ddot{e}_T$ . However, the integral in Eq. (3.9) gives us:

$$e_T = \frac{e_r}{2} (2(\tilde{p} + R)^2 + R^2) \quad (3.11)$$

The energy consumed by the processing before the re-transmission is denoted  $e_p$ . This energy is consumed by all nodes inside the zones. Therefore, the total energy  $E$  of a node transmitting data becomes:

$$E = e_p \rho z + \frac{e_r}{2} (2(\tilde{p} + R)^2 + R^2) \quad (3.12)$$

where  $\rho$  is the network nodes density per area unit.

### 3.2.5 Packet loss probability

Now, we can find the probability  $P_l$  of losing a transmitted packet from a source to a destination. Packet loss occurs when at least one of the zones between the source and destination does not include any node. Thus, we can formulate  $P_l$  as follows:

$$P_l = 1 - \left[ 1 - \left( \frac{A - z}{A} \right)^N \right]^n \quad (3.13)$$

where  $A$  is the area of overall network,  $z$  is the area of zone,  $N$  is the number of nodes in the network and  $n$  is the number of zones along the traffic.

*Proof.* See appendix A.6.  $\square$

### 3.2.6 System latency

The latency of sending a packet from a source to a destination has a significant impact on the performance of our protocol. When a packet reaches a current zone, all nodes in this zone initialise different random delay values  $t_i$  with  $0 < t_i < w$ , where  $w$  is the maximum delay. Let  $\alpha_i = \frac{t_i}{w}$ , hence, a node  $i$  chooses a random  $\alpha_i$  where  $0 < \alpha_i < 1$ . Therefore, a zone, containing  $\frac{z}{A}N = \rho z$  nodes, re-transmits a packet after a delay of  $\alpha_{min}w$ , where  $\alpha_{min} = \min\{\alpha_i\}$ . The probability that a zone transmits packet before a given delay  $vw$  can be written as follow:

$$\begin{aligned} F_{\alpha_{min}}(v) &= P(\alpha_{min} < v) \\ &= 1 - (1 - v)^{\rho z} \end{aligned} \quad (3.14)$$

where  $0 < v < 1$  and  $F_{\alpha_{min}}(v)$  is a Cumulative Distribution Function (CDF). So, we can calculate the Probability Density Function (PDF):

$$f(v) = \begin{cases} \rho z(1 - v)^{\rho z - 1} & \text{if } 0 < v < 1 \\ 0 & \text{otherwise} \end{cases} \quad (3.15)$$

Then, the time average  $T_{avg}$  needed to re-transmit packet by a zone is equals to:

$$\begin{aligned} T_{avg} &= w \int_0^1 \rho z(1 - v)^{\rho z - 1} v dv \\ &= w \left. \frac{-(1 - v)^{\rho z + 1}}{\rho z + 1} \right|_0^1 = w \frac{1}{\rho z + 1} \end{aligned} \quad (3.16)$$

## 3.3 Evaluation of our protocol

We compare the proposed LIBRO's efficiency to that of the state-of-the-art protocol DIR to assess its efficacy based on several criteria, which are listed in table 3.8. While both protocols use a fixed size header, LIBRO is free of loops, unlike DIR, which may occur. Because a node must know the geographical location of all its neighbours, nodes in DIR should be immobile. Nodes in LIBRO, on the other hand, can be mobile because a node does not need to know its location or the identities of its neighbours. Furthermore, DIR must perform some calculations before re-transmitting the packet, whereas the proposed LIBRO can do so immediately.

Various numerical results are performed to check the lifespan, packet loss probability,

and latency of the proposed protocol in order to thoroughly investigate its performance. We consider a WSN deployed over an area of  $200 \times 200 m^2$ , LIBRO zones of radius of  $3 m$ ,  $e_r = 0.5 nJ/m^2$ , the initial energy for each node is assumed to be  $10 J$ , maximum delay  $w = 0.1 ms$  and the elapsed time by DIR WSN node while calculating the distance from one of its neighbor is  $10 \mu s$ .

	LIBRO	DIR
Header	Fixed size header	Fixed size header
Loop	Free of loop	Loop may occurs
Mobility of nodes	Supports nodes mobility	Does not support nodes mobility
Neighbor nodes	Does not need any information about neighbor nodes	Always needs to know the geographical location of neighbors
Preprocessing before transmission	Transmits packet immediately	Calculates the distance of all its neighbors from the destination

FIGURE 3.8: Comparison table between LIBRO and DIR.

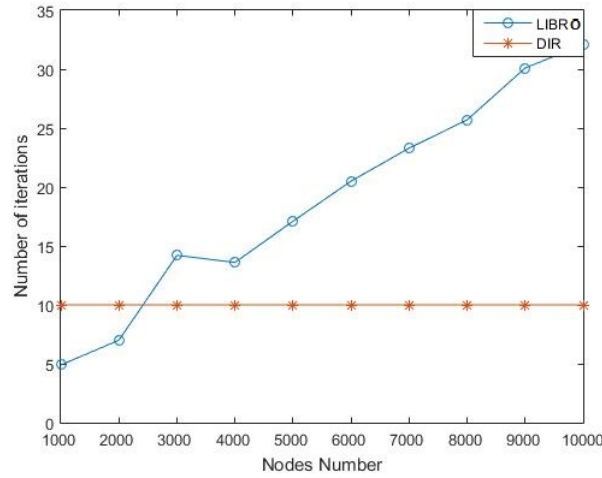


FIGURE 3.9: Lifespan comparison between LIBRO and DIR protocols

### 3.3.1 Energy consumption

Fig. 3.9 depicts the network's lifespan in terms of the number of existing nodes. In our case, the lifespan is defined as the number of iterations (a complete packet transmission between the source and the destination) that can be performed before a node dies. As shown in the graph, LIBRO's lifespan increases as the number of nodes increases, whereas DIR maintains the same performance. LIBRO outperforms DIR after 2000 nodes. The reason for this is that packets in DIR always follow the same trajectory between a source and a destination; thus, the same nodes are frequently used, whereas

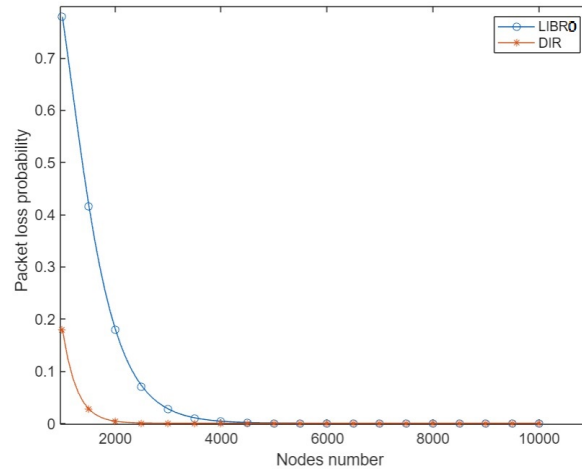


FIGURE 3.10: Packet loss probability between source and destination.

in LIBRO, each packet passes through a different trajectory, which distributes the transmission load over the nodes in LIBRO zones located in the selected trajectory.

### 3.3.2 Packet loss and latency

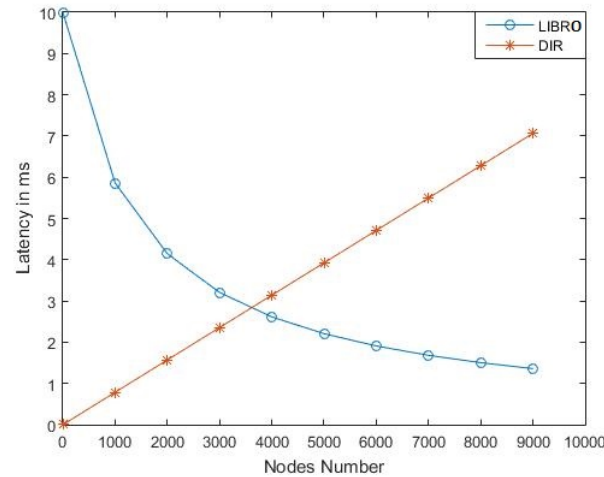


FIGURE 3.11: Latency of transmission of a packet from source to destination.

The probability of packet loss is shown in Fig. 3.10. LIBRO has a higher loss probability than DIR, especially when there are fewer nodes. This refers to the fact that, in comparison to DIR, each hop in LIBRO targets a small zone. This increases the likelihood of LIBRO having a hole (an empty zone). With a large number of nodes, LIBRO performs similarly to DIR, and with a very low packet loss probability, the possibility of having

an empty zone becomes negligible.

The latency is shown in Fig. 3.11. DIR is better than LIBRO for a small number of nodes; as a result, in LIBRO, the latency inside the zones increases as the number of nodes decreases Eq. (3.16). When there are a large number of nodes, there is a greater chance of selecting a small delay, which reduces latency. The latency in DIR grows with the number of nodes and the number of neighbour nodes. As a result, the time required by a node to calculate the nearest neighbour increases.

### 3.4 Conclusion

We proposed a new routing protocol called LIBRO. Rather than the traditional node routing, our protocol is based on zone routing using location information. When transmitting a packet, the node does not require any information about the identities of other nodes along the path to the destination. For high node density in the WSN, our protocol outperforms the well-known DIR. Furthermore, the proposed protocol can be thought of as a minimum requirement routing protocol. Also, the positive effect of increasing the number of nodes in the network, which leads to extended lifespan and reduced packet delays, clearly demonstrates scalability. In the following chapter, we will use LIBRO in an energy harvesting CRSN to improve its energy consumption and lifespan.

## Chapter 4

# Routing Optimization In Energy Harvesting Cognitive Radio Sensor Network

The energy consumption and connectivity are regarded as critical factors in determining the CRSN's efficiency. As a result, an efficient routing protocol is required to ensure reliable connectivity between network nodes while also significantly reducing energy consumption. Furthermore, energy harvesting in CRSN refers to compensating for the energy consumed by network nodes and extending network lifespan. As a result, we implement the Location Information-based Routing protocol (LIBRO) in the Cognitive Radio WSN (CRSN) with energy harvesting, this protocol is called CR-LIBRO. The lifespan of the CRSN is extended by solving an optimization problem while maintaining the lowest possible collision probability with PU and minimising data transmission latency. The simulation results validate the efficiency of our proposed method.

### 4.1 Introduction and Related works

In CRSN, the local physical data gathered by each node is transmitted to a Base Station (BS) [1] in a single or multi-hop manner [111]. CRSN consumes a large amount of wireless spectrum during the data collection phase, also known as the traffic discovery phase [112].

The authors of [113] propose a cognitive adaptive MAC (CAMAC) protocol that addresses the issue of power limitation in CRSNs by using on-demand Spectrum Sensing (SS), limiting the number of SS nodes by using a duty cycle, and significantly reducing

energy consumption. Another paper in [114] looks at the effect of packet size on energy efficiency in CRSN. The authors develop an optimization problem to determine the optimal packet size that ensures acceptable transmission efficiency and low energy consumption during packet transmission. To solve this optimization problem, the Sequential Quadratic Programming (SQP) method is used.

In [115], the authors address the delay-sensitive challenge in multimedia communication, and they present an open research issue for the realisation of energy-efficient and real-time transport in CRSN.

Data stored in a node in any subarea must be reachable at any time in order to ensure reliable communication among all nodes in the network. However, if the distance between the BS and a node is greater than the node's transmission range, we must configure a multi-hop routing protocol that achieves two main goals: the first is that all nodes are able to forward data to a BS in a multi-hop manner for distances greater than the node's transmission range. The second goal, is to improve the energy saving since the consumed energy is reduced by using multi-hop low range transmission instead of one large single-hop transmission.

Multi-hop routing protocols, such as Ad-hoc On-Demand Distance Vector (AODV), Dynamic Source Routing (DSR), Zone Routing Protocol (ZRP), Location-Aided Routing (LAR), Trajectory-Based Forwarding (TBF), and Distance Routing protocol (DIR), are widely studied in WSN to achieve the aforementioned goals. In the previous chapter, 3, we proposed LIBRO, a new routing protocol based on the Global Positioning System (GPS). In this protocol, a user wants to read data from a specific zone  $Z_0$ , regardless of the node ID delivering the data. As a result, an arbitrary node from zone  $Z_0$  is responsible for collecting data and transmitting it to the BS in a multi-hop manner, because data is transmitted between zones rather than moving between nodes. LIBRO requires less processing and memory, and WSN nodes can easily handle it.

#### 4.1.1 Related Work on Energy Harvesting

Energy Harvesting (EH) is a cost-effective technique for obtaining energy from renewable sources such as the sun, wind, heat, or radio frequency (RF). As a result, providing this technique to WSN results in the liberation of WSN nodes from energy constraints. Thus, EH may provide, under certain conditions, a perpetual energy source that energises the WSN nodes and eliminates the need to replace the nodes' batteries on a regular basis [38].

Improving energy efficiency is a critical factor in CRSN. The authors of [47] propose a



new scheme in which the SUs in the CR network harvest energy from two RF sources: the first is the PU transmission signal when the PU is active, and the second is the Secondary Relays (SRs) when they transmit their local detection results to the SU. The authors formulate a multi-objective optimization problem in this scheme in order to maximise the detection probability of PU using a cooperative sensing scenario, while minimising false alarms and thus increasing spectrum efficiency.

The authors of [116] develop an infinite-horizon discounted Markov decision process to minimise the outage probability of an EH in CRSN based on battery energy, channel fading, and harvested energy. The EH technique used in this scheme aims to improve the system's self-sustainability. In this Markov decision process, two policies are used: the sensing-transmission (ST) policy and the efficient transmission (ET) policy, which is used only when the SNR is sufficiently high.

Authors of [117] investigate the optimal SS interval in EH in CR networks. The authors demonstrate the existence of a trade-off between the throughput of the SU and the interference to the PU in this work. This trade-off is formulated as a Partially Observable Markov Decision Process (POMDP) and a Markov Decision Process (MDP) optimization problem, in order to find the optimal SS while maximising the network's throughput.

The authors of [118] divide the CR network into two groups. The SU nodes in the second group harvest energy from the PU RF transmission signal when the SU nodes in the first one perform cooperative SS. If the PU is detected to be active (in transmission mode), the SU in Group 2 helps the PU to transmit its data by amplifying the PU's RF signal, while the SU in Group 1 harvests energy. The goal of this work is to determine the optimal number of SU in the two groups, the cooperative power gain for the SU in Group 2 during PU signal amplifying, the SU transmission power, and the sensing parameters (sensing period and sensing threshold).

#### 4.1.2 Chapter's Contribution

We propose a CRSN with EH capability; our network's transmission is based on the LIBRO routing protocol, which requires little memory and processing power. A multi-objective optimization problem is formulated to find the optimum value of CR node transmission range in order to minimise collision probability with PU, minimise energy consumption (i.e. maximising lifetime even to reach the self-sustainable system), minimise network transmission latency, and finally maximise network throughput. LIBRO is simple to be implemented in CRSN, and it uses very low amount of energy. As a result, it allows for the efficient integration of EH techniques into the network. This chapter's main contributions can be summarised as follows:

- In our EH CRSN, we use the LIBRO routing protocol. This protocol is very easy to implement in CRSN because of its simplicity; it requires little memory because the node does not require any prior knowledge of its neighbours' IDs, nor does it require any information about the network topology. Nodes with a low processing capability are required to find a route between source and destination. Furthermore, the efficiency of this protocol is unaffected by node mobility. The use of this protocol in our EH CRSN reduces the amount of energy used during the traffic search process. AODV routing protocol, on the other hand, floods the network with RREQ messages during route discovery.
- The energy constraints in the network may be eliminated by incorporating the EH technique into the CRSN. A CRSN node, even if deployed in a harsh environment, does not require battery replacement on a regular basis. Our research aims to identify network parameters that will allow the network to be powered by a renewable energy source indefinitely, i.e. make energy consumption less than the energy harvested by the EH units installed in all network nodes.
- Finding a balance between the following factors: PU collision probability, SU energy consumption during SU activities (SS, route discovery, data transmission), SU delivery packet latency, and network throughput. In this work, we create a self-sustaining network that meets all network criteria. We do so by formulating a multi-objective optimization problem and then finding a solution that meets all required criterias.

## 4.2 System Model

We consider a CRSN of area  $A$  composed of  $N$  nodes (which may be mobile nodes) and one or more BS distributed along the network. These nodes are distributed at random over a large geographical zone. As in the previous chapter's system model, a radio wake-up scheme is used to ensure that a node that wants to transmit a message to the next zone can wake-up all nodes in its coverage transmission region, and that they are listening and ready to receive and forward messages. A GPS module or other tools are provided to the nodes in order for them to be aware of their locations. They are powered by a battery and have an initial energy  $e_{int}$ .

### 4.2.1 CR Network Model

Because all nodes in our network have CR capability, each node can sense the spectrum and transmit data to the BS opportunistically using a licensed channel when the PU is

Notations List	
Notation	Definition
$N$	Total node numbers in the network
$A$	Area of geographical zone of the network
$\rho$	Network nodes density
$R$	Radius of LIBRO zone
$z$	Area of LIBRO zone
$N_{PU}$	Number of PU in the network geographical zone
$N_{CH}$	Number of licensed channels
$w$	Maximum chosen delay
$t$	The current chosen delay
$n$	Number of steps between the Target and the Destination zones
$P(H_1)$	Probability of PU to be active
$P(H_0)$	Probability of PU to be idle
$P_{md}$	Probability of miss detection of the PU
$P_{fa}$	Probability of false alarm in CR model
$T_{SW}$	Channel switch delay
$p$	Maximum fixed length of LIBRO step
$\tilde{p}$	Calculated step length where $\tilde{p} < p$
$\ddot{r}_t$	Exact transmission range of network node
$r_t$	Approximated transmission range of network node
$\mathbf{r}$	seonde approximation of transmission range of network node
$\ddot{e}_T$	Exact consumed transmission energy
$e_T$	Approximated consumed transmission energy
$e_r$	Energy density at the receiver antenna
$e_p$	Consumed energy during node processing
$e_s$	consumed energy during SS
$E$	Total consumed energy of node without SS
$E_T$	Total consumed energy of node with SS
$P_l$	Probability of data loss
$\epsilon$	Quantitative harvested energy by a single node

TABLE 4.1: The notations.

idle. We consider one channel to be high priority, and all nodes use it if it is free. If the higher priority channel is retained by a licensed PU, the SU must search for another free channel to use [119]. We assume that a channel that is sensed to be idle remains idle throughout the transmission phase of an SU. We also take into account the network geographical zone contains several PUs distributed randomly in the network (Fig. 4.1), Let  $N_{PU}$  is the number of PUs.

We also have  $N_{CH}$  licensed channels dedicated for the PUs, which the PU can access and use for transmission-reception purposes; as previously stated, one of these channels has a higher priority to be used by the SU.

Both PU and SU are assumed to use a time-slotted synchronous communication protocol with time slot length  $T$ . The PU model is based on a two-state discrete Markov process

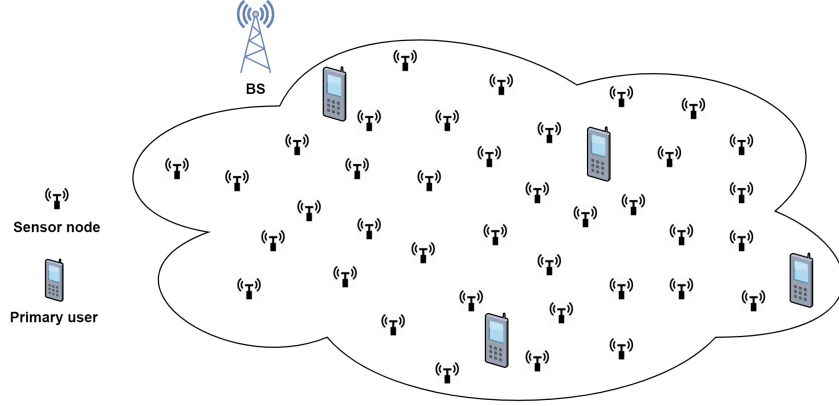


FIGURE 4.1: Cognitive Radio Sensor Network.

[120–123], in which the PU has two states, one of which is active when the PU is in transmission or receive mode, and the other is idle. The prior probability of the PU being active or idle is  $P(H_1)$  and  $P(H_0)$  respectively, where  $H_1$  and  $H_0$  are the hypothesis of PU to be active or idle respectively. In our case, the SU consists of the WSN nodes and the BS. As a result, all nodes in this network can sense the CR spectrum. Because the sensing process is imperfect in practise, when a Sensor Node (SN) performs a sensing action, it may make miss-detection when the PU is present but the SU misses it. Another error occurs when the SU detects PU activity while the PU is not present; in this case, the state is known as false alarm. As a result, the probabilities of miss-detection and false alarm are  $P_{md}$  and  $P_{fa}$ , respectively. Accordingly, the probability of colliding with the PU during an SU transmission phase can be expressed as follows:

$$c = P(H_1)P_{md} \quad (4.1)$$

where the PU state is assumed to be unchanged during a complete time slot  $T$ .

The probability where the SU believes there is an active PU in its SS zone is as follow:

$$b = P(H_1)(1 - P_{md}) + P(H_0)P_{fa} \quad (4.2)$$

In our system model, if the SU wants to transmit data, but it detects the presence of a PU, then it switches to another free channel where, this action consumes time, and the time needed to switch from one channel to another is  $T_{SW}$ .

#### 4.2.2 CR-LIBRO routing model

In the previous chapter, we proposed a routing protocol called LIBRO, which uses GPS to route a data message from a source to a destination. We include this protocol in our proposed scheme, which means that each sensor node in our network should have a GPS unit to be aware of its geographical location.

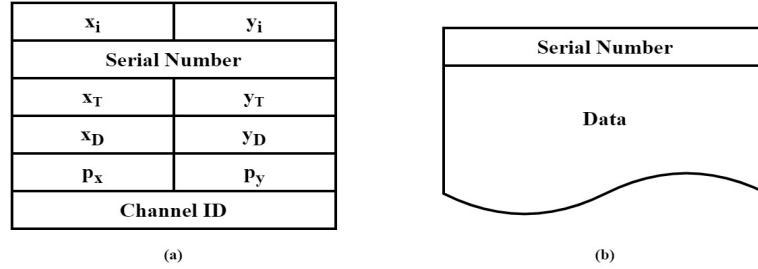


FIGURE 4.2: (a) Control message through CCC, (b) Physical data message.

Fig. 3.2 shows an example of collecting data from a target zone, then transmitting the message in a multi-hop fashion from the target to the gateway by jumping between zones rather than specific nodes. The process of CR-LIBRO is depicted in Fig. 4.3, the network is linked to a BS (shown as a gateway in the figure); the BS is accessible to users via the internet. When a user wants data from a specific zone of radius  $R$ , it sends a request to the gateway via an internet connection. The gateway broadcasts a control message containing its geographical location, the location of the target zone, the location of the next zone, the steps length, and a serial number via single hop transmission via the Common Control Channel (CCC). As a result, all nodes within the target zone start a random count down counter  $t$  between 0 and  $w$ , where  $w$  is the maximum available delay. The first node within the target zone that exceeds its delay period performs SS and then selects a free channel. Then it constructs a control packet containing a serial number, its geographical location, BS location, step length, as well as the location of the next zone and free data channel ID. Following that, this node broadcasts a control message toward the next zone in the direction of the BS via CCC, followed by a data message containing the collected physical data with the same serial number via the data channel indicated in the control message. Simultaneously, nodes within the same transmitter zone receive that control message; as a result, they stop counting down timers and drop the message because someone else did their job. Following that, only nodes in the next zone respond to this message, tuning their receiver to the proposed channel determined in the control message, then receiving the data message, and the process repeats until the message reaches the BS, at which point all nodes in the destination zone drop the message, and the BS sends data to the end user. Assuming free-space transmission, the

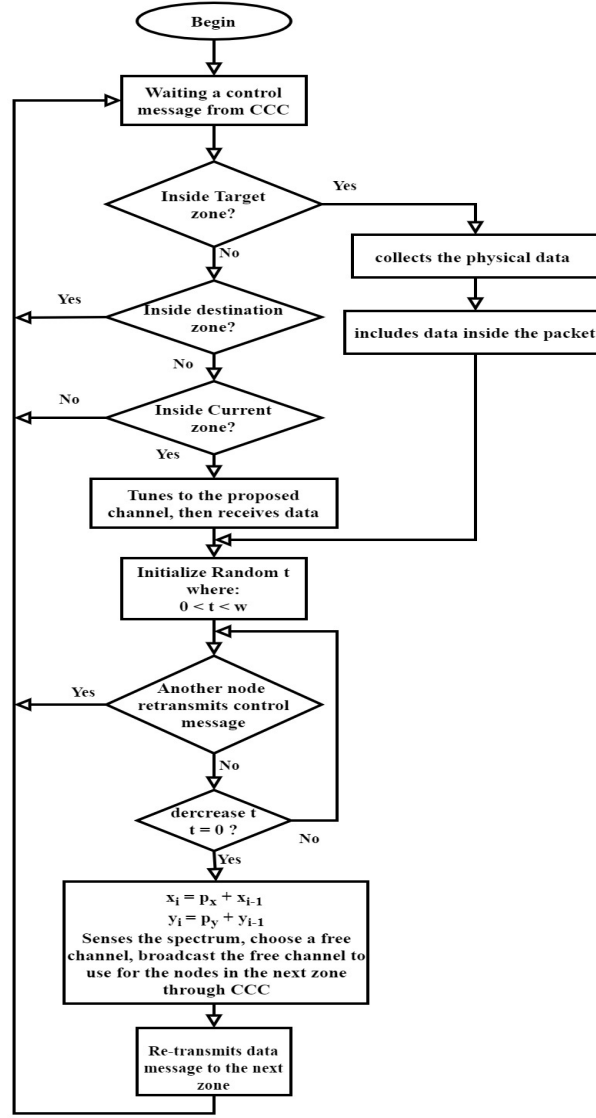


FIGURE 4.3: CR-LIBRO flow chart.

average amount of energy required to send a data packet between two consecutive zones is  $E_T$  which is calculated as follow:

$$E_T = E + e_s \quad (4.3)$$

where  $E$  is the consumed energy by a zone calculated in the previous chapter Eq. (3.12),  $e_s$  is the energy consumed by SS.

### 4.2.3 Energy harvesting routing model

The EH features are integrated into the network nodes to improve energy efficiency. This allows the network to compensate for the energy loss that occurs as a result of its operations. Each node is capable of harvesting energy from the environment, such as light, wind, heat, and radio frequency signals. The collected energy is then used to energise its internal battery. Each node collects energy for each time slot. We assume the internal battery is never fully charged and must be restored at all times. Furthermore, a Zone Activity Period (ZAP) is defined as all activities performed between the time a zone receives a message and the time the zone re-transmits the message; we assume that each time slot is equal to one ZAP. The quantum harvested energy by a node during one time slot or ZAP is defined as  $\epsilon$ . During a ZAP, the total amount of harvested energy is:

$$h_Z = \epsilon \rho z \quad (4.4)$$

where  $\rho$  is the networks' nodes density,  $z$  is the area of the LIBRO zone.

## 4.3 Problem Formulation

The primary goal of this section is to determine the best value for the SU transmission range  $r_t$  and the CR-LIBRO zones area  $z$ . This results in a good system performance in terms of minimising the collision probability with the PU, improving the network's energy efficiency by reducing the energy consumption of each node in order to provide the system with inexhaustible energy thanks to the EH technique, minimising communication latency, and maximising network throughput.

### 4.3.1 Collision Probability

Collision with PU is a very critical for CRSN; Thus, we aim to reduce this collision probability  $P_c$  by keeping its value below some maximum permissible limit  $P_c^*$ . In our network, a collision with the PU occurs when an active PU is within the transmission range  $r_t$  of the SU during the SU transmission phase. By referring to Eq. (3.11) in the previous chapter, we can deduce the average transmission range  $\bar{r}_t$  which consumes  $e_T$  joules:

$$\bar{r}_t = \sqrt{\frac{2(\tilde{p} + R)^2 + R^2}{2}} \quad (4.5)$$

In order to avoid the overlapping between CR-LIBRO zones along a traffic, the CR-LIBRO step  $\tilde{p}$  must be twice the radius of CR-LIBRO zone  $R$ . Giving  $\tilde{p} \geq 2R$ , we can approximate the value of  $\bar{r}_t$  as follow:

$$\bar{r}_t = \sqrt{\frac{2(\tilde{p} + R)^2 + R^2}{2}} \approx \tilde{p} + R = \mathfrak{r} \quad (4.6)$$

*Proof.* See appendix A.5. □

Knowing that a SU is in the transmission phase, the likelihood that the SU will collide with a PU is as follows:

$$P_c = \left[ 1 - \left( 1 - \frac{\pi \mathfrak{r}^2}{A} \right)^{N_{PU} P(H_1)} \right] P_{md} \quad (4.7)$$

where  $\mathfrak{r}$  is the SU transmission range,  $\left( 1 - \frac{\pi \mathfrak{r}^2}{A} \right)^{N_{PU} P(H_1)}$  is the probability of a transmission zone does not contain any active PU.

*Proof.* See appendix A.7. □

However,  $P_c$  must be always under some maximum permissible value  $P_c^*$ , hence, we have:

$$P_c = \left[ 1 - \left( 1 - \frac{\pi \mathfrak{r}^2}{A} \right)^{N_{PU} P(H_1)} \right] P_{md} \leq P_c^* \quad (4.8)$$

our goal is to reduce this collision probability.

### 4.3.2 Energy Consumption and harvesting

The CRSN's primary goal is to improve energy efficiency. Because of the harsh geographic zone in which the CRSN nodes are deployed, replacing the battery is not always a simple task; thus, we aim to reduce the network's energy consumption in order to extend the network's lifespan.

During the SS, transmission, and processing phases, a node consumes the most of its available energy. The energy consumed within one zone during message transmission is given by:

$$E_T = e_p \rho z + e_r \mathfrak{r}^2 + e_s \quad (4.9)$$

where  $e_s$  is the energy consumed during the SS phase,  $e_p$  is the energy consumed during the processing phase,  $\rho$  is network density measured by nodes per area unit, and  $z$  is the zone area. Within a zone, one node performs SS and data transmission during



one packet transmission, consuming energy during SS and data transmission; all nodes within the zone consume energy during the processing phase.

To maintain a continuous source of energy, the used energy must be less than the harvested energy for a whole cycle of data message routing, i.e. the time required to transport a data message from a target zone to the BS:

$$E_T \leq \epsilon \frac{D}{\tilde{p}} \rho z = \epsilon \frac{D}{\mathbf{r} - R} \rho z \quad (4.10)$$

where  $\epsilon$  is the harvested energy occurs by a single node during one ZAP,  $D$  is the expected distance between the BS and the target zone,  $\tilde{p}$  is the distance between two consecutive LIBRO zones. The term  $\frac{D}{\tilde{p}}$  is the number of hops between the BS and the target zone, each hop need ZAP seconds in order to be completely achieved.

*Proof.* See appendix A.8. □

#### 4.3.3 Transmission packet latency

In this work, we define data latency as the total amount of time required for physical data to be transmitted from a target zone to the BS. The time required for a zone to re-transmit a data message after receiving it is  $T_{avg}$  as given in the original LIBRO Eq. (3.16).

CRSN activities such as SS process, receive a control message, and switch to another channel consume time, and they cannot be ignored in this case. Therefor, the total time needed by a physical data to be transmitted from a target zone to the BS is:

$$T_p = \left[ w \frac{1}{\rho z + 1} + T_{SW} (P(H_1)(1 - P_{md}) + P(H_0)P_{fa}) + t_s + t_r \right] \frac{D}{\mathbf{r} - R} \quad (4.11)$$

where  $T_{SW}$  is the switched time to another free channel if the first channel is not idle, this term is multiplied by the probability where the SU believes that the PU is active on this channel (the channel with the higher priority).  $t_s$  is the SS time,  $t_r$  is the time needed to receive the control and the data message. As we can see, each ZAP needs  $T_p \frac{\mathbf{r}-R}{D}$  seconds in order to be achieved. In order to improve our system,  $T_p$  must be reduced.

#### 4.3.4 Packet Delivery Ratio

Packet Delivery Ratio refers to the ratio of received packets to sent packets. The data message is deemed lost in this work if it is transmitted to an empty zone or a zone

with empty power nodes; all other data loss is ignored. The packet delivery ratio is determined as:

$$\lambda = \left[ 1 - \left( \frac{A-z}{A} \right)^N \right]^{\frac{D}{\tau-R}} \quad (4.12)$$

we aim to minimize this ratio.

#### 4.3.5 Loop-free condition

In order to ensure that our protocol guarantees loop-free routing, overlapping CR-LIBRO zones must be avoided. Therefore, the radius  $R$  of the CR-LIBRO zone should not exceed half the step  $\tilde{p}$  value. Thus, we can determine the maximum area  $z$  of the CR-LIBRO zone. Using the equation Eq. (4.6) we have:

$$\tau = \tilde{p} + R \geq 3R \implies \pi\tau^2 \geq 9\pi R^2 \implies \pi\tau^2 \geq 9z \quad (4.13)$$

This condition must be respected during the problem optimization.

#### 4.3.6 Optimization problem formulation and resolution

To recapitulate, by grouping all previous individual optimization problems we obtained a multi-objective optimization problem defined as:

$$\min_{\tau} : P_c = \left[ 1 - \left( 1 - \frac{\pi\tau^2}{A} \right)^{N_{PU}P(H_1)} \right] P_{md} \quad (4.14a)$$

$$\min_{\tau, z} : E_T = e_p \rho z + e_r \tau^2 + e_s \quad (4.14b)$$

$$\min_{\tau, z} : T_p = \left[ w \frac{1}{\rho z + 1} + T_{SW} (P(H_1)(1 - P_{md}) + P(H_0)P_{fa}) + t_s + t_r \right] \frac{D}{\tau - R} \quad (4.14c)$$

$$\max_{\tau, z} : \lambda = \left[ 1 - \left( \frac{A-z}{A} \right)^N \right]^{\frac{D}{\tau-R}} \quad (4.14d)$$

$$s.t. : P_c \leq P_c^* \quad (4.14e)$$

$$E_T \leq \epsilon \frac{D}{\tau - R} \rho z \quad (4.14f)$$

$$\pi\tau^2 \geq \alpha z \quad (4.14g)$$

where  $\alpha \geq 9$  is a real constant used in order to ensure that the area of a CR-LIBRO zone is less than 9 times the area of transmission range as shown in equation Eq. (4.13). To tackle our multi-objective optimization issue, we employ the lexicographic method.

Rather of providing weights, the lexicographic method imposes preferences by arranging the goal functions according to their value or significance. The following optimization problems are thus solved one at a time.

Therefor, we begin resolving the collision probability in the equation (4.14a). We consider that  $P_c^*$  is the maximum permissible collision probability with PU, hence, we can calculate an upper boundary of  $\tau$ :

$$\tau \leq \sqrt{\frac{A}{\pi} \left[ 1 - \sqrt{(N_{PU} P(H_1))} \sqrt{1 - \frac{P_c^*}{P_{md}}} \right]} \quad (4.15)$$

*Proof.* See appendix A.9. □

For the energy consumption equation (4.14b), in order to restore the overall consumed energy, which extends the network lifespan, we must ensure that:

$$e_r \tau^2 + e_s + e_p \rho z \leq \epsilon \frac{D}{\tau - R} \rho z \quad (4.16)$$

By resolving this above equation we get:

$$z \geq \frac{1}{\rho} \left( \frac{e_r \tau^2 + e_s}{\epsilon \frac{D}{\tau - R} - e_p} \right) \quad (4.17)$$

$z$  is strictly positive, wherefore, the denominator in the above formula  $\epsilon \frac{D}{\tau - R} - e_p$  can be pushed to a strictly positive value which gives us:

$$\tau \leq \frac{\epsilon D}{e_p} + R \quad (4.18)$$

Combining all inequalities of  $\tau$  in equations (4.15) and (4.18) we get:

$$\tau \leq \min \left( \frac{\epsilon D}{e_p} + R, \sqrt{\frac{A}{\pi} \left[ 1 - \sqrt{(N_{PU} P(H_1))} \sqrt{1 - \frac{P_c^*}{P_{md}}} \right]} \right) \quad (4.19)$$

As well as, combining all inequalities of  $\tau$  in equations (4.14g) and (4.17) we get:

$$\frac{1}{\rho} \left( \frac{e_r \tau^2 + e_s}{\epsilon \frac{D}{\tau - R} - e_p} \right) \leq z \leq \frac{\pi \tau^2}{\alpha} \quad (4.20)$$

Now, by looking to the last two equations (4.14c) and (4.14d), we can notice that, if  $\tau$  and  $z$  increase, then,  $T_p$  decreases and  $\lambda$  increases (this is actually our goal). Thus, in order to get the optimum values of  $T_p$  and  $\lambda$ , we must choose the maximum allowed

values of  $\mathbf{r}$  and  $z$ . Hence, the solution of our problem is given by:

$$\hat{\mathbf{r}} = \min \left( \frac{\epsilon D}{e_p} + R, \sqrt{\frac{A}{\pi} \left[ 1 - N_{PU} P(H_1) \sqrt{1 - \frac{P_c^*}{P_{md}}} \right]} \right) \quad (4.21a)$$

$$\hat{z} = \frac{\pi \hat{\mathbf{r}}^2}{\alpha} \quad (4.21b)$$

#### 4.4 Numerical results.

We consider a geographical area of 100 metres width and 500 metres length where SU and PU nodes are installed in our experiment. Both miss-detection and false alarm have a probability of 0.1. Each PU has a 0.4 chance of being active. The transmission parameter density  $et = 0.001watt/m^2$ , the energy consumed during SS  $es = 0.1watt$ , and the processing energy  $ep = 0.01$ . The other characteristic will be revealed later.

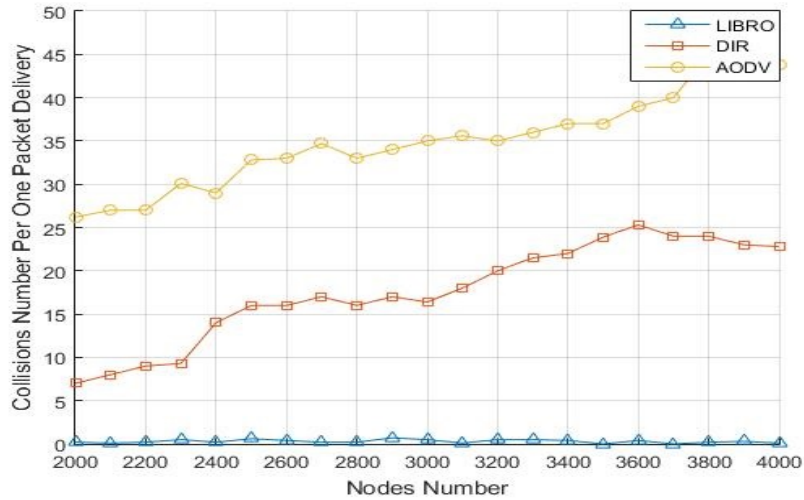


FIGURE 4.4: The number of collision between SU and PU occurred during the overall packet delivery period with different node number.

##### 4.4.1 Comparison with AODV and DIR

Several simulations are provided to evaluate the performance of our system. The AODV and DIR protocols are compared with our protocol. The number of collisions for our CR-LIBRO, DIR, and AODV protocols during the complete delivery packet period is shown in Fig. 4.4. The number of PUs is 10, and we will assume we only have one licensed channel for the sake of simplicity. As can be seen, our protocol has the lowest

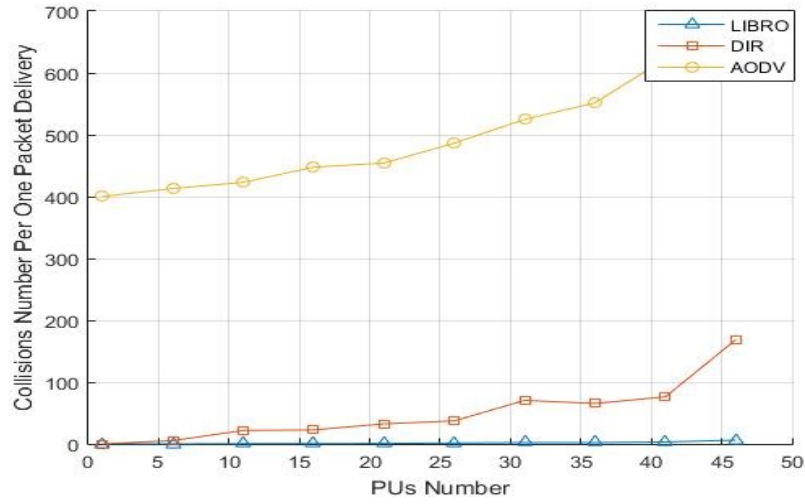


FIGURE 4.5: The number of collision between SU and PU occurred during the overall packet delivery period with different PU number.

amount of collisions; as a result, during the re-transmission phase, one node in each zone makes a transmission. In DIR, a node asks location data from its neighbours after each re-transmission; during this phase, all neighbours must send their locations, potentially increasing the collision chance. The RREQ message is re-broadcast by all nodes in the network during the route request phase in AODV, which is why the AODV has a higher rate of collisions during the packet delivery phase. We witness the similar effect in Fig. 4.5 by adjusting the number of PU and keeping the number of nodes at 2000.

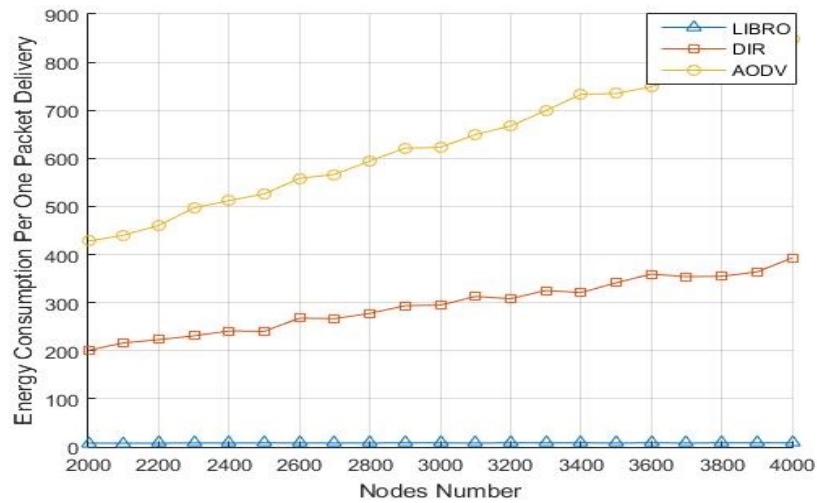


FIGURE 4.6: Consumed energy by the overall network during the overall packet delivery period with different nodes number.

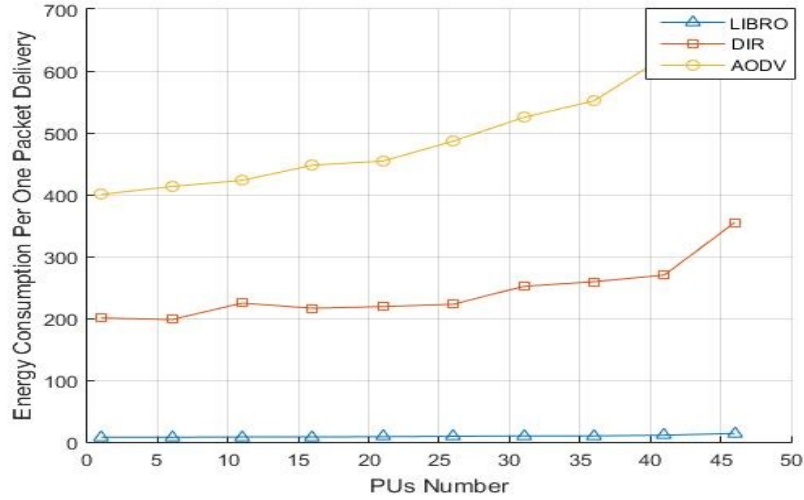


FIGURE 4.7: Consumed energy by the overall network during the overall packet delivery period with different PU number.

Our protocol CR-LIBRO outperforms DIR and AODV in term of energy consumption as shown in Fig. 4.6 and Fig. 4.7. Because the number of transmissions in CR-LIBRO is equal to the number of hops, it consumes less energy than DIR and AODV. As demonstrated in Fig. 4.7, increasing the PU number has a comparable effect on energy consumption. When a transmission collides with a PU, the SU must re-transmit its packet, which uses additional energy.

The difference in packet delivery time is seen in Fig. 4.8. The simulation results reveal

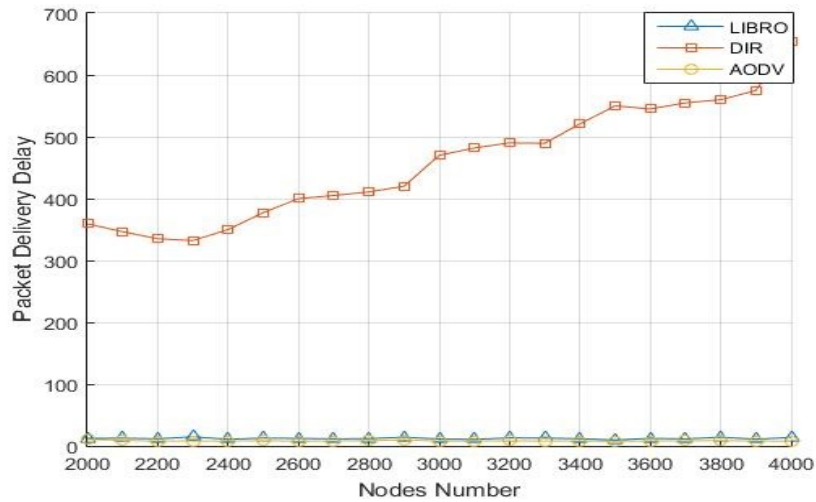


FIGURE 4.8: Delay between generating the data message by a target node, and receiving it by the BS, with different nodes number.

that our protocol and AODV are similar in that they both require the same number of hops for a packet to reach its destination. However, because DIR collects position data

from current transmitter neighbours at each hop, the packet takes longer to reach its target. The number of nodes in the network does not affect the latency in CR-LIBRO as it does in AODV, whereas DIR takes longer to transmit a packet as the number of nodes increases.

As shown in Fig. 4.9, increasing the number of PUs in the network causes more DIR

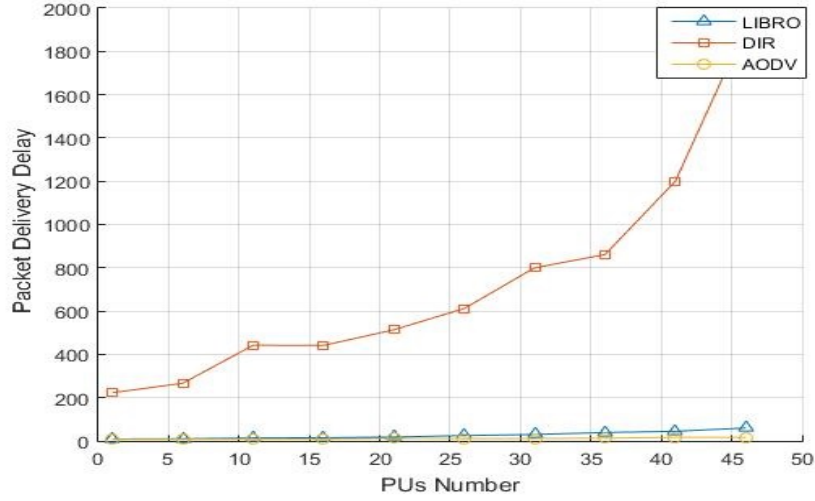


FIGURE 4.9: Delay between generating the data message by a target node, and receiving it by the BS, with different PU number.

collisions, which causes packet delivery delays. Such effect is due to the quite number of collecting data about position from neighbors, but it has a considerably less impact on the CR-LIBRO and AODV protocols.

Finally, the table in Fig. 4.11 highlights the key differences between the three protocols:

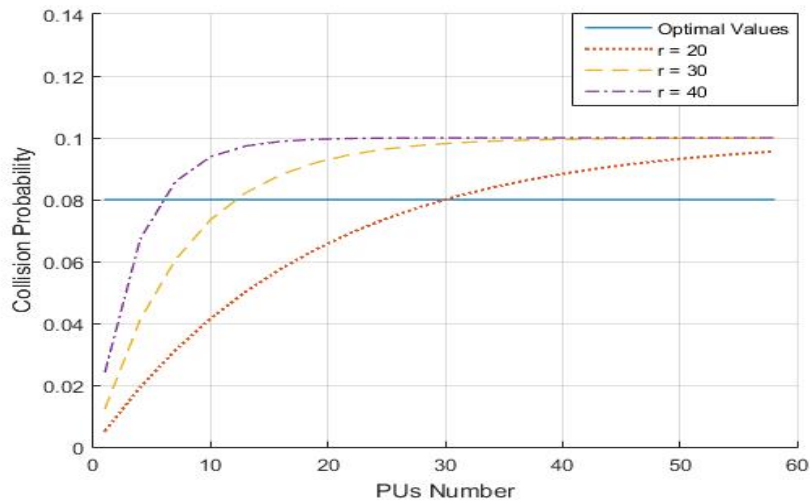


FIGURE 4.10: Comparison of collision probability with different values of transmission range and the optimal value.

	LIBRO	DIR	AODV
Loop	Free of loop	Loop may occurs	Free of loop
Neighbor nodes	Does not need any information about neighbor nodes	Always need to know the geographical location of neighbor	Always need to know the ID of its neighbor nodes
Preprocessing before transmission	Transmits packet immediately	Calculates the distance of all its neighbors from the destination	Route request is needed before transmission
Mobility nodes	Support nodes mobility without any addition processing	Energy consumption increases with mobility	Energy consumption increases with mobility
Scalability	Latency decreases and PDR increases by increasing the nodes number	Increasing the nodes number has a negative effect on the latency, collision and energy consumption	Increasing the nodes number has a negative effect on the collision and energy consumption

FIGURE 4.11: Comparison table between LIBRO, DIR and AODV.

CR-LIBRO, DIR, and AODV. This table demonstrates our system's superiority in terms of loop, mobility, memory, processing, and scalability.

#### 4.4.2 Evaluation of optimization results

We compare our optimal values with arbitrary values in order to evaluate our optimization outcomes, as shown in Fig. 4.10, 4.12, 4.14 and 4.13. In our network, the maximum collision probability allowed is 0.08. By increasing the number of PUs in the network, the collision probability increases with different transmission ranges, as shown in Fig. 4.10. However, by selecting the best coverage range for each PUs number, we may have the best coverage range while keeping the collision probability within acceptable bounds..

Our system with the optimal value can save energy and lead us to a system with an endless energy supply by harvesting energy from the environment and restoring all consumed energy, as shown in Fig. 4.12. We must keep our energy use below the quantity we can gather in order to keep our system energy-perpetual. The packet delivery delay is illustrated in Fig. 4.13. As the number of nodes increases, the delay decreases, as shown in the equation Eq. (4.11). A large number of nodes in a zone causes the zone to choose a minimal delay before re-transmission. As we can see, by choosing optimal value of coverage transmission range  $r_t$ , and the ratio between zone and range  $\alpha = \frac{z}{\pi r_t^2}$ , we can accelerate our system. The packet delivery ratio, as shown in Fig. 4.14, increases as the number of nodes in the network grows, implying that the probability of finding an empty zone lowers as the number of nodes and hence the density of the network grows. The optimal parameters allow us to get the highest possible throughput.



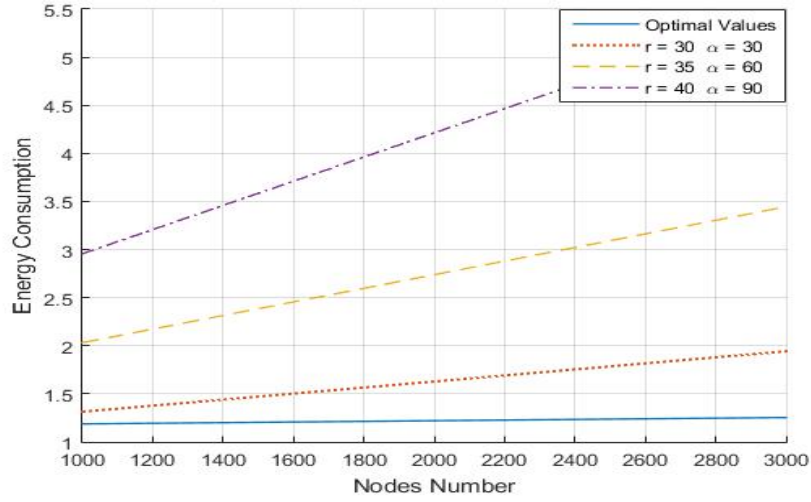


FIGURE 4.12: Comparison of energy consumption with different values of transmission range, zone area and the optimal value.

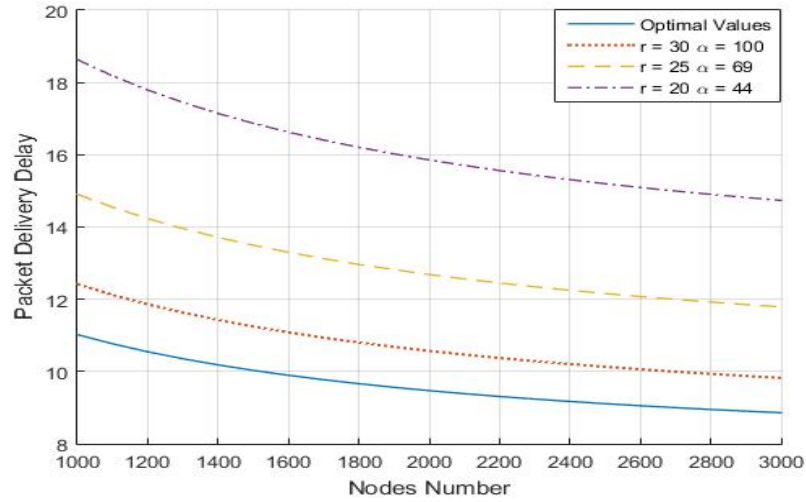


FIGURE 4.13: Comparison of packet delivery delay with different values of transmission range, zone area and the optimal value.

## 4.5 Conclusion

In this chapter, we looked at a CRSN that harvests energy with the LIBRO protocol to increase energy efficiency by recovering all spent energy and spectrum efficiency by reducing collisions with the PU. Furthermore, we address an optimization problem to get the maximum packet delivery ratio with the least amount of latency. The simulation results demonstrate the success of our system by comparing it to the AODV and DIR protocols, demonstrating that our system, which uses CR-LIBRO, outperforms other

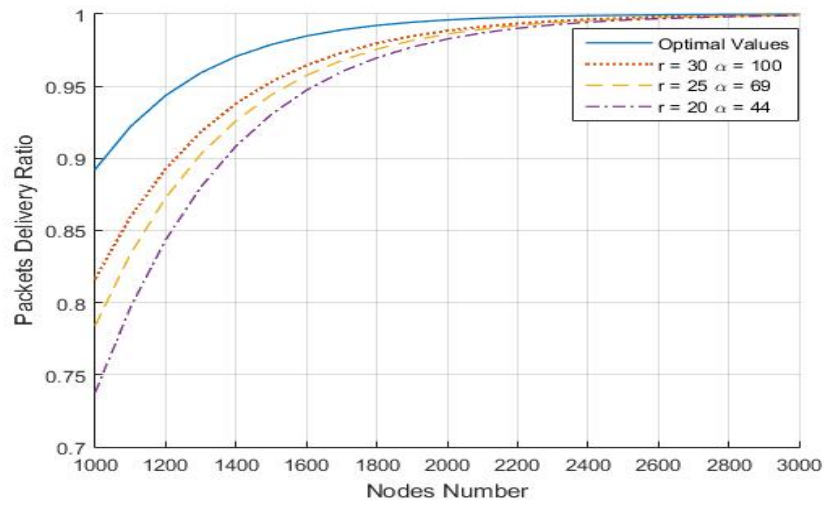


FIGURE 4.14: Comparison of packet delivery ratio with different values of transmission range, zone area and the optimal value.

systems that use the AODV and DIR protocols. Furthermore, mathematical calculations indicate the feasibility of constructing a system with perpetual energy in which all consumed energy may be regenerated using the harvesting process.

## Chapter 5

# General Conclusion

In this thesis, we have reviewed the characteristics of WSNs, as well as the issue of energy consumption in these networks, with its significant impact on the lifespan of these networks.

In the first chapter, we reviewed in general the principle of WSNs, their characteristics and challenges. We also talked about the issue of clustering and routing for data transmission in a simple and guaranteed manner. Then we touched on the issue of the scarcity of spectrum resources, and reviewed the CR principle and its importance in improving the efficiency of the spectrum. After that, we talked about the possibility of integrating the characteristics of CR into the WSNs. Finally, we presented the importance of using the energy harvesting mechanism in these networks, which prolongs the life of the network and turns it into an environmentally friendly network with green energy power sources.

In the chapter 2, we took into account the value of the energy consumed during the performance of the SS process in CRSN, in order to reach the optimal number of clusters in these networks, which reduces energy consumption. We also developed mathematical equations to calculate the average number of packets sent through each node, which led us to calculate the extent of the data latency.

In Chapter 3, we have proposed a new routing protocol called LIBRO that can be used in mobile WSNs, because of its simplicity in implementation in those networks, so that it does not flood the whole network with unnecessary messages during route discovery process, and it does not require complex processing and also does not need a large memory space.

In Chapter 4, we took advantage of LIBRO's ability to include it in the CRSN in order to obtain the CR-LIBRO protocol, which gave us an impetus to use energy harvesting techniques in an attempt to completely free the network from energy constraints. We

compared our methods with other well-known protocols such as AODV and DIR, and our system showed a clear superiority over its peers.

## 5.1 Contributions

In the introduction to each chapter, we have mentioned our contributions, however we will rephrase them all in this section, and we will mention additional contributions:

- For the chapter 2, we strive for the optimal number of CRSN clusters that take into account several parameters, such as: energy efficiency related to transmissions and sensing, and detection of CRSN cluster false alarm probabilities.
- To the best of our knowledge this is the first time the average number of packets relayed by each CH is obtained in a closed form. So that we can calculate the average number of packets sent by any cluster head within the network.
- Thanks to the equation for the average number of transmitted packets, the collision probability and the optimal CRSN cluster response time are derived, considering the reduction of data redundancy thanks to IDA, thus we can save a great deal of energy.
- The numerical results are performed to show the efficiency of choosing the optimum number of clusters, the relevant collision rate and response time, compared to networks where IDA is not used.
- In addition, the equation for the average number of transmitted packets has conducted us to a new equation for calculating the greatest common divisor, because we were able to write the first equation in two different ways, which led us to that result.
- In Chapter 3, we proposed a new routing protocol called LIBRO. This distributed protocol does not need infrastructure, and does not require high processing capacity or long processing time. Thus, it is superior to its peers, such as AODV and DIR. Nodes in networks that support this protocol do not need to store any information related to neighboring nodes such as their identities or geographical locations, so this protocol uses a small amount of memory space.
- Thanks to the processing speed and performance, LIBRO reduces the power consumption, as it relies on the GPS system to reduce the network overflow of unhelpful messages in route discovery stage. In addition, the packet header of transmitted message is small compared to other protocols, which saves more power consumption.

- Unlike analogues of the protocols that support the GPS system, LIBRO distributes the load to all nodes located within the LIBRO zones along the way. Every time a source sends data to its destination, this data will take different paths within the same zone route, because each time there will be a different node that will precede all of its peers in replying to the message and sending it towards its destination.
- Because it does not need any information about the identities and locations of the intermediate nodes between the user or the BS and the target area, this protocol works very efficiently in mobile networks, as it is not affected at all by the movement of nodes within the network.
- In chapter 4, We have integrated our LIBRO protocol into EH CRSN which effectively reduces power consumption during the traffic search process. Unlike the AODV routing protocol, which floods the network with RREQ messages during the route discovery process. The CR-LIBRO sends its messages to specific areas within the network.
- By adding EH technology in CRSN, the power constraint in the network is alleviated. CRSN nodes may be deployed in a harsh environment where periodic battery replacement is a difficult process. Our study aims to find some parameters of the network, in order to provide the network with a permanent source of green energy that is environmentally friendly.
- Reducing the total energy consumption in all network activities from the energy harvested by the EH unit integrated in all network nodes, may lead to providing the network by an eternal power source, and opens up options for us in front of a large number of new algorithms and features that we can add to the network that we were not able to previously used due to power limitations.
- Finding a trade-off between: probability of collision with PU, SU energy consumption during SU different activities (spectrum sensing, route discovery, data transmission), latency of SU delivery packets, and network throughput. In this work, we create a self-powering sustainable network while respecting all network standards, all done by formulating a multi-objective optimization problem, and then finding its solution that links the above criteria.

## 5.2 Future works

In future work, regarding the optimal clustering in the WSN that we reviewed in Chapter 2, we will consider the initial multi-PU, and we will incorporate other properties in

calculating the optimal value of the number of clusters within the network. The number of clusters has a significant role in reducing energy consumption, as it reduces the need for many spectral resources and reduces the interference in transmission within the network itself.

As for the issue of data routing, which we reviewed in the third and fourth chapters, our proposed protocol LIBRO suffers from a problem, which is sending data to an area completely free of any node (empty LIBRO zone), and therefore that message will not reach anyone and will be lost. Therefore, we will propose multiple solutions to confront this problem, making room for interested researchers to find more effective solutions. We can pave the way for more research:

- LIBRO is an on demand WSN routing protocol, where the BS specifies the area from which the user wants to collect data. When using the event driven pattern in wireless sensor networks, all sensors must be aware of the location of the BS. One of the solutions that we can suggest is that the BS periodically broadcasts a signal containing its geographical location over the entire network, in order to give the hand to all events to find their way to the BS. It is interesting to find the optimal period for the BS location signaling, specially in a mobile WSN where the speed of WSN nodes has an important impact on that period.
- The 5G technology opens the door to new kinds of routing in CRSN. Where, by using directive antenna for transmission introduced in a set of specific CHs with extra power source and 5G capabilities, the routing path that a message takes can wrap around an active PU's transmission range. Thus, the angle and the transmission range play a critical role in determining the routing path of each message in the CRSN.
- Using 5G directive transmission, the WSN nodes can determine their location. By replacing a number of beacon in the network, these beacons sweep a message along the network containing the angle of transmission and the beacon's ID. Each node receives a message, it stores the angle received from each beacon, the set of angle becomes the coordinates of each node. Simulations may depict the precision of geo-location.
- Clustering is a vital technology used in CRSN. However, due to the limited node resources, introducing all 5G features into the nodes is not an easy task. Alternatively, the nodes can be provided with directional antennas. Thus, after nodes are deployed in a geographic area, each node will take a random direction according to the orientation of its antenna. We can now install CHs with additional resources to broadcast their clustering offer across the network, allowing all nodes directed

to it to join its cluster. Each node will join the CH to which it is directed, and thus, will be Angular Clustering. This idea can be developed and the best location and number of cluster heads must be optimized.

## Chapter 6

# List Of Publications

- Mortada, M. R., Makhoul, A., Harb, H., Laiymani, D. and Barhamgi, M. (2019). EK-means: A new clustering approach for datasets classification in sensor networks. *Ad Hoc Networks*, 84, 158-169.
- Mortada, M. R., Makhoul, A., Abou Jaoude, C., Harb, H. and Laiymani, D. (2019, June). A distributed processing technique for sensor data applied to underwater sensor networks. In *2019 15th International Wireless Communications and Mobile Computing Conference (IWCMC)* (pp. 979-984). IEEE.
- Mortada, M. R., Nasser, A., Mansour, A. and Ya, K. C. (2019, September). Novel sensing mechanism for full-duplex secondary users in cognitive radio. In *2019 27th European Signal Processing Conference (EUSIPCO)* (pp. 1-5). IEEE.
- Mortada, M. R., Nasser, A., Mansour, A. and Yao, K. C. (2021, September). LIBRO: A Location Information Based Routing Protocol for Multi-Hop WSN Applications. In *2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall)* (pp. 1-6). IEEE.
- Mortada, M. R., Nasser, A., Mansour, A. and Yao, K. C. (2021). In-Network Data Aggregation for Ad Hoc Clustered Cognitive Radio Wireless Sensor Network. *Sensors*, 21(20), 6741.
- Mortada, M. R., Nasser, A., Mansour, A. and Yao, K. C. LIBRO in Energy Harvesting Cognitive Radio Wireless Sensor Network. *IEEE Sensors Journal*, submitted



## Appendix A

## Appendix A

### A.1 Proof of Eq. (2.20)

Assuming our CRSN contains  $P$  equal physical zones as shown in Fig.2.5. Each zone contains an equal number of CHs collecting similar physical data packets. We can divide our network into  $k_p$  physical rings, where the smallest ring is divided into 3 physical zones. Hence, the area of one physical zone is:

$$A_p = \frac{\pi}{3} \left( \frac{R}{k_p} \right)^2 \quad (\text{A.1})$$

The number of the physical zone  $P$  in the overall network is:

$$P = \frac{\pi R^2}{\frac{\pi}{3} \left( \frac{R}{k_p} \right)^2} = 3k_p^2 \quad (\text{A.2})$$

where  $R$  is the radius of overall CRSN. Similarly to Eq. (2.1), the area of  $i_{th}$  physical ring is:

$$P_{A_i} = \pi \left( \frac{R}{k_p} \right)^2 (2i - 1) \quad (\text{A.3})$$

where  $i \in [1, k_p]$ . The number  $m_{p_j}$  of physical zones in physical ring  $j$  is equal to:

$$m_{p_j} = \frac{P_{A_i}}{A_p} = 3(2i - 1) \quad (\text{A.4})$$

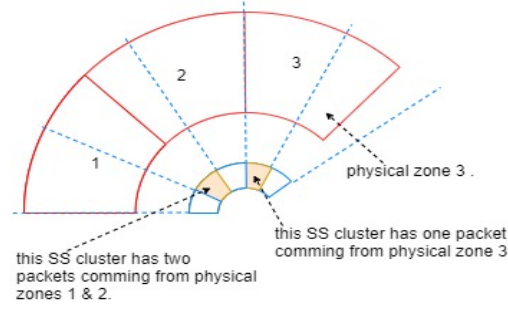


FIGURE A.1: The forwarding packets coming from the physical ring  $j$  by the CHs in spectrum ring  $i$

As shown in Fig.A.1, CHs in the SS ring  $i$  forward packets coming from physical zones located in the physical ring  $j$ . Obviously, in this example, we have CHs which forward one packet and others forward two packets coming from the same physical zone. In general, in the same SS ring we have a set of CHs which forward  $n$  packets and other which forward  $n + 1$  packets coming from the same physical ring. The average  $av_{i,j}$  of the number of forwarding packets coming from a physical ring  $j$  by the CHs in SS ring  $i$  is calculated as:

$$av_{i,j} = \frac{1}{m_i} \sum_{t=1}^{m_i} \left( \left\lceil \frac{m_{p_j}}{m_i} t \right\rceil - \left\lfloor \frac{m_{p_j}}{m_i} (t-1) \right\rfloor \right) \quad (\text{A.5})$$

The sum of average  $av_i$  of the number of forwarding packets coming from all physical zones by the CHs in spectrum ring  $i$  is:

$$av_i = \sum_{j=\lceil \frac{k_p}{k} i \rceil}^{k_p} \frac{1}{m_i} \sum_{t=1}^{m_i} \left( \left\lceil \frac{m_{p_j}}{m_i} t \right\rceil - \left\lfloor \frac{m_{p_j}}{m_i} (t-1) \right\rfloor \right) \quad (\text{A.6})$$

where  $j = \lceil \frac{k_p}{k} i \rceil$  is the index of the physical ring containing the current SS ring  $i$ .

Finally, in average, a CH located in the ring  $i$  in our CRSN, spends  $\tau_{T_i}$  to forwarding packets toward its next CH located in the ring  $i + 1$  in the way to the BS:

$$\tau_{T_i} = \frac{T_P}{m_i} \sum_{j=\lceil \frac{k_p}{k} i \rceil}^{k_p} \sum_{t=1}^{m_i} \left( \left\lceil \frac{m_{p_j}}{m_i} t \right\rceil - \left\lfloor \frac{m_{p_j}}{m_i} (t-1) \right\rfloor \right) \quad (\text{A.7})$$

## A.2 Proof of Eq. (2.21)

Let  $a = m_{p_j}$  and  $b = m_i$ , then:

$$\sum_{i=1}^b \left( \left\lceil \frac{a}{b} i \right\rceil - \left\lfloor \frac{a}{b} (i-1) \right\rfloor \right) = a + b - (a \wedge b) \quad (\text{A.8})$$

therefore, we have:

$$(a \wedge b) = a + b - \sum_{i=1}^b \left( \left\lceil \frac{a}{b} i \right\rceil - \left\lfloor \frac{a}{b} (i-1) \right\rfloor \right) \quad (\text{A.9})$$

= Let  $(a \wedge b) = c$ , then  $a = cM$  and  $b = cN$  with  $M \wedge N = 1$ .

1. If  $a = b$ , then  $(a \wedge a) = a$  and:

$$a + b - \sum_{i=1}^b \left( \left\lceil \frac{a}{b} i \right\rceil - \left\lfloor \frac{a}{b} (i-1) \right\rfloor \right) \quad (\text{A.10})$$

$$= a + b - \sum_{i=1}^b (\lceil i \rceil - \lfloor i-1 \rfloor) \quad (\text{A.11})$$

$$= a + b - \sum_{i=1}^b 1 = a \quad (\text{A.12})$$

2. If  $a > b$ , then we get:

$$a + b - \sum_{i=1}^b \left( \left\lceil \frac{M}{N} i \right\rceil - \left\lfloor \frac{M}{N} (i-1) \right\rfloor \right) \quad (\text{A.13})$$

$$= b - \sum_{i=1}^{b-1} \left( \left\lceil \frac{M}{N} i \right\rceil - \left\lfloor \frac{M}{N} i \right\rfloor \right) \quad (\text{A.14})$$

as we can see, if  $\frac{M}{N} i \in \mathbb{N}$  then  $\lceil \frac{M}{N} i \rceil - \lfloor \frac{M}{N} i \rfloor = 0$

In this case, we have  $i \in S = \{N, 2N, \dots, (c-1)N\}$ , then we get:

$$b - \sum_{i=1}^{b-1} \left( \left\lceil \frac{M}{N} i \right\rceil - \left\lfloor \frac{M}{N} i \right\rfloor \right) \quad (\text{A.15})$$

$$= b - \left( \sum_{i=1}^{b-1} 1 - \text{card}(S) \right) \quad (\text{A.16})$$

$$= b - b + c = c \quad (\text{A.17})$$

3. If  $b > a$ : it can be proved similarly to the case of  $a > b$

### A.3 Proof of Eq. (2.26)

According to Eq. (A.2), If  $k_p = 1$ , that means we have just 3 distinct physical zones, in this case:

$$N_{p_i} = \frac{1}{m_i} \sum_{j=1}^1 \left( m_{p_1} + m_i - (m_{p_1} \wedge m_i) \right) \quad (\text{A.18})$$

where, according to Eq. (A.4),  $m_{p_1} = 3$ . We get:

$$N_{p_i} = \frac{1}{m_i} \sum_{j=1}^1 m_i = 1 \quad (\text{A.19})$$

which means all CHs in the overall network relay just one packet. So, let us consider the case where  $k_p \geq 2$ . Hence, to demonstrate that  $N_{p_1} > N_{p_i}$ ,  $\forall i > 1$ , we should prove that:

$$k_p^2 - \frac{1}{m_i} \sum_{j=\alpha}^{k_p} \left( m_{p_j} + m_i - (m_{p_j} \wedge m_i) \right) > 0 \quad (\text{A.20})$$

where  $\alpha = \lceil \frac{k_p}{k} i \rceil$ . Having:

$$\frac{1}{m_i} \sum_{j=\alpha}^{k_p} \left( m_{p_j} + m_i \right) > \frac{1}{m_i} \sum_{j=\alpha}^{k_p} \left( m_{p_j} + m_i - (m_{p_j} \wedge m_i) \right) \quad (\text{A.21})$$

Inequality (A.20) can be bounded by inequality (A.21). We should prove:

$$k_p^2 - \frac{1}{m_i} \sum_{j=\alpha}^{k_p} \left( m_{p_j} + m_i \right) > 0 \quad (\text{A.22})$$

Hence:

$$\begin{aligned}
& k_p^2 - \frac{1}{m_i} \sum_{j=\alpha}^{k_p} (m_{p_j} + m_i) \\
&= k_p^2 - \frac{1}{3(2i-1)} \sum_{j=\alpha}^{k_p} (3(2j-1) + 3(2i-1)) \\
&= k_p^2 - \frac{1}{(2i-1)} \sum_{j=\alpha}^{k_p} ((2j-1) + (2i-1)) \\
&= k_p^2 - \frac{1}{(2i-1)} \left( k_p^2 + k_p(2i-1) + (1-\alpha)(2i+\alpha-2) \right) \\
&= k_p^2(2i-2) - k_p(2i-1) + (\alpha-1)(2i+\alpha-2) \tag{A.23}
\end{aligned}$$

In the last equation,  $(\alpha-1)(2i+\alpha-2) \geq 0$  since  $1 \leq \alpha \leq k_p$  and  $i > 1$ . Finally, we have to prove that  $k_p^2(2i-2) - k_p(2i-1) > 0$ . For this, we solve the following equation:

$$k_p^2(2i-2) - k_p(2i-1) = 0 \tag{A.24}$$

Solving the equation above, we get:

$$\begin{aligned}
& k_p = 0, \text{ unacceptable since } k_p \geq 2 \\
& k_p = \frac{2i-1}{2i-2} < \frac{3}{2}, \forall i > 0, \text{ unacceptable since } k_p \geq 2
\end{aligned}$$

Therefore, the last equation has the same sign as  $k_p^2$ . Hence, Eq. (A.23) is always positive. Finally, the inequality (A.20) is proved.

## A.4 Proof of Eq.(3.10)

In order to demonstrate that the ratio of average consumed energy  $\ddot{e}_T$  in Eq.(3.7) and the average consumed energy  $e_T$  in Eq.(3.9), we must suppose the worst case where the radius  $R$  of LIBRO zones is equals to  $\frac{\bar{p}}{2}$ . Obviously, the biggest difference between the exact transmission range  $\ddot{r}_t$  and the approximate transmission range  $r_t$  occurs when  $M$  is located on the edge of previous LIBRO zone  $Z_{i-1}$  as shown in Fig.A.2. Let  $\delta = \frac{r_t}{\ddot{r}_t}$  the ratio of a exact and approximate transmission range of individual node in zone  $Z_i$  existing on the edge of this zone. The maximum value of  $\delta$  can be found by calculating the derivative  $\frac{d\delta}{d\theta}$  as follow:

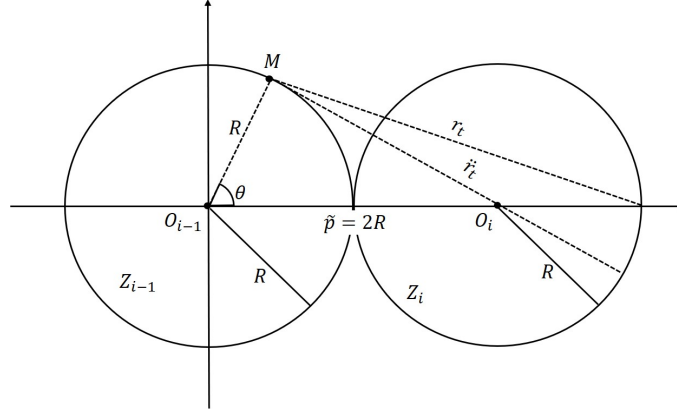


FIGURE A.2: Caption

$$\begin{aligned} \frac{d\delta}{d\theta} &= \left( \frac{\sqrt{(3R)^2 + R^2 - 2(3R)R \cos \theta}}{\sqrt{(2R)^2 + R^2 - 2(2R)R \cos \theta} + R} \right)' = \left( \frac{\sqrt{10 - 6 \cos \theta}}{\sqrt{5 - 4 \cos \theta} + 1} \right)' \\ &= \frac{(3\sqrt{-4 \cos \theta + 5} - 5) \sin \theta}{\sqrt{-6 \cos \theta + 10} \sqrt{-4 \cos \theta + 5} (\sqrt{-4 \cos \theta + 5} + 1)^2} \end{aligned} \quad (\text{A.25})$$

In order to find the maximum of  $\delta$  with respect to  $\theta$  we must resolve  $\frac{d\delta}{d\theta} = 0$ , thus we must resolve the following equation:

$$(3\sqrt{-4 \cos \theta + 5} - 5) \sin \theta = 0 \quad (\text{A.26})$$

However, for  $\theta = 0$ ,  $M$  is on the  $x$  axis, hence, we have  $\ddot{r}_t = r_t$ , thus,  $\theta$  must be different from 0, then,  $\sin \theta \neq 0$ . Therefore, the Eq.(A.26) becomes:

$$(3\sqrt{-4 \cos \theta + 5} - 5) = 0 \quad (\text{A.27})$$

By resolving the equation Eq.(A.27) we got:

$$\cos \theta = \frac{5}{9} \quad (\text{A.28})$$

We substitute the value of  $\cos \theta$  of Eq.(A.28) in the equation of  $\delta$ , then we got:

$$\delta = \frac{r_t}{\ddot{r}_t} \approx 97\% \quad (\text{A.29})$$

which is the maximum percentage of difference between  $r_t$  and  $\ddot{r}_t$ . Finally, we calculate the ratio  $\frac{e_T}{\ddot{e}_T}$  as following:

$$\frac{e_T}{\ddot{e}_T} \geq \delta^2 \approx 94\% \quad (\text{A.30})$$

Thus,  $\ddot{e}_T$  can be approximated by  $e_T$ .

## A.5 Proof of Eq.(4.6)

Giving  $\tilde{p} \geq 2R$ , by choosing the lowest value  $\tilde{p} = 2R$  then substitute  $\tilde{p}$  in the equation of  $r_t$  in Eq.(4.6) we got:

$$\sqrt{\frac{2(2R+R)^2 + R^2}{2}} \approx 2R + R \quad (\text{A.31})$$

Hence, by dividing the two terms we got:

$$\frac{2R + R}{\sqrt{\frac{2(2R+R)^2 + R^2}{2}}} = \frac{3R}{\sqrt{\frac{19R^2}{2}}} \approx 97\% \quad (\text{A.32})$$

Thus we have:

$$\frac{\tilde{p} + R}{r_t} \geq 97\% \quad (\text{A.33})$$

Therefor,  $r_t = \tilde{p} + R$  is a good approximation.

## A.6 Proof of Eq.(3.13)

To ensure that the message reaches from LIBRO zone  $X$  to zone  $Y$ , each zone along the route between  $X$  and  $Y$  must have at least one node. Therefore, if at least one zone is empty of nodes, the message between  $X$  and  $Y$  is lost.

To calculate the probability that each zone along the path between  $X$  and  $Y$  contains at least one node each, we must convert the problem into an enumeration problem. So we divide the area  $A$  of the whole network by the area  $z$  of the LIBRO zone. Thus, we have divided the network into  $M = \frac{A}{z}$  zones.

The problem has become a classic enumeration issue. We first assume that the nodes are completely similar, and thus calculate the number of cases  $c_t$  in which  $N$  nodes can be distributed within  $M$  regions:

$$c_t = C_{N+M-1}^N = \frac{(N+M-1)!}{N!(M-1)!} \quad (\text{A.34})$$

Then we calculate all the cases in which the nodes are distributed over the  $M$  zones, provided that  $n$  regions among them contain at least one node each:

$$c_n = C_{N-n+M-1}^{N-n} = \frac{(N-n+M-1)!}{(N-n)!(M-1)!} \quad (\text{A.35})$$

Next, we calculate the loss probability  $\tilde{P}_l$  of losing a message sent between  $X$  and  $Y$ :

$$\tilde{P}_l = 1 - \left( \frac{(N-n+M-1)!}{(N-n)!(M-1)!} \right) \left( \frac{N!(M-1)!}{(N+M-1)!} \right) = 1 - \frac{N!(N-n+M-1)!}{(N-n)!(N+M-1)!} \quad (\text{A.36})$$

Replacing  $M$  by  $\frac{A}{z}$  we get:

$$\tilde{P}_l = 1 - \frac{N!(N - n + \frac{A}{z} - 1)!}{(N - n)!(N + \frac{A}{z} - 1)!} \quad (\text{A.37})$$

The simulation showed a significant difference in the results in calculating the loss probability of losing the message between the simulation and the above equation Eq. (A.37). The reason for this is that the computer program, while distributing the nodes in the network, it does so uniformly, and therefore we will not find the case in which the nodes are gathered in one area only, for example.

Thus, we conclude that because of the uniform distribution of nodes in the network, the number of nodes in a particular LIBRO zone is an independent event as it does not affect the number of nodes within other zones. Thus, we calculate the probability of the message being lost as follows: First, the probability  $P_e$  of a LIBRO zone to be empty is formulated as follow:

$$P_e = \left( \frac{A - z}{A} \right)^N \quad (\text{A.38})$$

Thus, the probability  $\bar{P}_l$  of having one node at least in each LIBRO zone along the whole traffic is writing as follow:

$$\bar{P}_l = \left[ 1 - \left( \frac{A - z}{A} \right)^N \right]^n \quad (\text{A.39})$$

We were able to write the last equation because of the independence of events that describe the distribution of nodes within zones in the network. Therefor, the packet loss occurs when at least one zone along the traffic is empty, this packet loss probability  $P_l$  is deduced as follow:

$$P_l = 1 - \left[ 1 - \left( \frac{A - z}{A} \right)^N \right]^n \quad (\text{A.40})$$

## A.7 Proof of Eq.(4.7)

The number of active PU in the network is:

$$N_{PU}^{Active} = N_{PU}P(H_1) \quad (\text{A.41})$$

The collision occurs with the PU when at least one active PU fall within the transmission range of a SU node. Thus, the probability where a current transmission range doesn't



contains any PU is as follow:

$$P_0 = \left(1 - \frac{\pi \mathfrak{r}^2}{A}\right)^{N_{PU} P(H_1)} \quad (\text{A.42})$$

Then, the probability of having at least one active PU within the SU transmission range, as well as, the SU miss detect this PU. Therefore, a collision may occurs and the probability of collision becomes:

$$P_c = \left[1 - \left(1 - \frac{\pi \mathfrak{r}^2}{A}\right)^{N_{PU} P(H_1)}\right] P_{md} \quad (\text{A.43})$$

## A.8 Proof of Eq.(4.10)

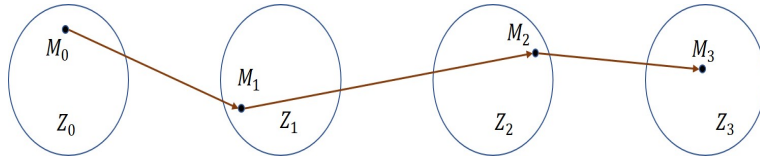


FIGURE A.3: Packet transmission cycle

Figure Fig.(A.3) depicts the data transmission cycle from the source  $M_0$  to the destination  $M_3$ . After a data message packet goes from zone  $Z_0$ , this zone consumes an amount of energy  $e_T$ . Thus, this zone must recycles all of its energy at the end of the cycle (from  $M_0$  to  $M_3$ ). In this case we have 3 hops, the number of hops is equals to  $\frac{D}{\tilde{p}}$ . At each hop a zone collects  $e_h$  joules:

$$e_h = \epsilon \rho z \quad (\text{A.44})$$

where  $\rho z$  is the number of nodes inside a zone. At the end of data transmission cycle from source to destination we have  $n_h$  hops:

$$n_h = \frac{D}{\tilde{p}} = \frac{D}{\mathfrak{r} - R} \quad (\text{A.45})$$

Thus, at the end of whole data transmission a zone collects:

$$E_h = \epsilon \frac{D}{\mathfrak{r} - R} \rho z \quad (\text{A.46})$$

## A.9 Proof of Eq.(4.15)

By looking to the equation Eq.(4.8), this inequality is resolved as follow:

$$\begin{aligned}
\left[1 - \left(1 - \frac{\pi \mathfrak{r}^2}{A}\right)^{N_{PU}P(H_1)}\right] P_{md} &\leq P_c^* \\
1 - \left(1 - \frac{\pi \mathfrak{r}^2}{A}\right)^{N_{PU}P(H_1)} &\leq \frac{P_c^*}{P_{md}} \\
\left(1 - \frac{\pi \mathfrak{r}^2}{A}\right)^{N_{PU}P(H_1)} &\leq 1 - \frac{P_c^*}{P_{md}} \\
\left(1 - \frac{\pi \mathfrak{r}^2}{A}\right) &\geq \sqrt[N_{PU}P(H_1)]{1 - \frac{P_c^*}{P_{md}}} \\
1 - \frac{\pi \mathfrak{r}^2}{A} &\geq \sqrt[N_{PU}P(H_1)]{1 - \frac{P_c^*}{P_{md}}} \\
\frac{\pi \mathfrak{r}^2}{A} &\leq 1 - \sqrt[N_{PU}P(H_1)]{1 - \frac{P_c^*}{P_{md}}} \\
\mathfrak{r}^2 &\leq \frac{A}{\pi} \left[1 - \sqrt[N_{PU}P(H_1)]{1 - \frac{P_c^*}{P_{md}}}\right] \\
\mathfrak{r} &\leq \sqrt{\frac{A}{\pi} \left[1 - \sqrt[N_{PU}P(H_1)]{1 - \frac{P_c^*}{P_{md}}}\right]}
\end{aligned}$$

## Appendix B

## Appendix B

### B.1 Simulation of double integral in equations Eq.(3.7) and Eq.(3.9)

#### B.1.1 Java program:

```
public class MainProg {

    public static double deploy(double R){
        return 2*R*Math.random()-R;
    }

    public static double distance(double x0,double y0,double x1,double y1){
        return Math.sqrt((x0-x1)*(x0-x1)+(y0-y1)*(y0-y1));
    }

    public static void main(String[] args) { // TODO Auto-generated method stub
        double R=10;
        double P=100;
        double exactInt=0;
        double approxInt=0;
        int iteration=10000000;
        double x=0;
        double y=0;
        double number=0;
```

```
for(int i=0;i<iteration;i++){
    x=deploy(R);
    y=deploy(R);

    if(distance(x, y, 0, 0)<=R){
        number++; approxInt+=Math.pow(distance(x, y, P+R, 0),2);
        exactInt+=Math.pow(distance(x, y, P, 0)+R,2);
    }
}

System.out.println("the approximated result is "+approxInt/number);
System.out.println("the exact result is "+exactInt/number);
System.out.println("the ratio is "+approxInt/exactInt);

double calculatedValue=Math.pow(P+R,2)+R*R/2;
System.out.println("the approximative calculated value is "+calculatedValue);

}

}
```

### **B.1.2 Results:**

the approximated result is: 12150.1180  
the exact result is: 12152.6187  
the ratio is: 0.9997  
the approximative calculated value is: 12150.0

# Bibliography

- [1] Mohammad Abdul Matin and MM Islam. Overview of wireless sensor network. *Wireless sensor networks-technology and protocols*, pages 1–3, 2012.
- [2] Gang Zhou, John A Stankovic, and Sang H Son. Crowded spectrum in wireless sensor networks. *IEEE EmNets*, 6, 2006.
- [3] Ekram Hossain and Vijay K Bhargava. *Cognitive wireless communication networks*. Springer Science & Business Media, 2007.
- [4] Ayaz Ahmad, Sadiq Ahmad, Mubashir Husain Rehmani, and Naveed Ul Hassan. A survey on radio resource allocation in cognitive radio sensor networks. *IEEE Communications Surveys & Tutorials*, 17(2):888–917, 2015.
- [5] Ozgur B Akan, Osman B Karli, and Ozgur Ergul. Cognitive radio sensor networks. *IEEE network*, 23(4):34–40, 2009.
- [6] Mohammad Abdul Matin and MM Islam. Overview of wireless sensor network. *Wireless sensor networks-technology and protocols*, pages 1–3, 2012.
- [7] Sang H Kang. Energy optimization in cluster-based routing protocols for large-area wireless sensor networks. *Symmetry*, 11(1):1–37, 2019.
- [8] Jamal N Al-Karaki and Ahmed E Kamal. Routing techniques in wireless sensor networks: a survey. *IEEE wireless communications*, 11(6):6–28, 2004.
- [9] Eiko Yoneki and Jean Bacon. A survey of wireless sensor network technologies. *technical report, University of Cambridge*, 2005.
- [10] Marcos Augusto M Vieira, Claudionor N Coelho, DC jr da Silva, and José Monteiro da Mata. Survey on wireless sensor network devices. In *EFTA 2003. 2003 IEEE Conference on Emerging Technologies and Factory Automation.*, volume 1, pages 537–544. IEEE, 2003.
- [11] Daisuke Takaishi, Hiroki Nishiyama, Nei Kato, and Ryu Miura. Toward energy efficient big data gathering in densely distributed sensor networks. *IEEE transactions on emerging topics in computing*, 2(3):388–397, 2014.

- [12] Yang Yang, Rick S Blum, and Brian M Sadler. A distributed and energy-efficient framework for neyman-pearson detection of fluctuating signals in large-scale sensor networks. *IEEE Journal on Selected Areas in Communications*, 28(7):1149–1158, 2010.
- [13] Ameer Ahmed Abbasi and Mohamed Younis. A survey on clustering algorithms for wireless sensor networks. *Computer communications*, 30(14-15):2826–2841, 2007.
- [14] Mian Ahmad Jan, Priyadarsi Nanda, and Xiangjian He. Energy evaluation model for an improved centralized clustering hierarchical algorithm in wsn. In *International Conference on wired/wireless internet communication*, pages 154–167. Springer, 2013.
- [15] SR Boselin Prabhu, S Sophia, PD Manivannan, and S Nithya. A research on decentralized clustering algorithms for dense wireless sensor networks. *International Journal of Computer Applications*, 57(20), 2012.
- [16] Wendi Beth Heinzelman. *Application-specific protocol architectures for wireless networks*. PhD thesis, Massachusetts Institute of Technology, 2000.
- [17] D Dugaev, S Zinov, E Siemens, and V Shuvalov. A survey and performance evaluation of ad-hoc multi-hop routing protocols for static outdoor networks. In *2015 International Siberian Conference on Control and Communications (SIBCON)*, pages 1–11. IEEE, 2015.
- [18] D Dugaev, S Zinov, and E Siemens. A survey of multi-hop routing schemes in wireless networks applied for the smartlighting scenario//v international science conference. *Technologies and equipment for information measurement Tomsk, Russia*, 2014.
- [19] Ian D Chakeres and Elizabeth M Belding-Royer. Aodv routing protocol implementation design. In *24th International Conference on Distributed Computing Systems Workshops, 2004. Proceedings.*, pages 698–703. IEEE, 2004.
- [20] Yingpeng Sang, Hong Shen, Yasushi Inoguchi, Yasuo Tan, and Naixue Xiong. Secure data aggregation in wireless sensor networks: A survey. In *2006 Seventh International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT'06)*, pages 315–320. IEEE, 2006.
- [21] L Krishnamachari, Deborah Estrin, and Stephen Wicker. The impact of data aggregation in wireless sensor networks. In *Proceedings 22nd international conference on distributed computing systems workshops*, pages 575–578. IEEE, 2002.

- [22] Marwa Ibrahim, Hassan Harb, Ali Mansour, Abbass Nasser, and Christophe Osswald. All-in-one: Toward hybrid data collection and energy saving mechanism in sensing-based iot applications. *Peer-to-Peer Networking and Applications*, 14(3): 1154–1173, 2021.
- [23] Marwa Ibrahim, Hassan Harb, Abbass Nasser, Ali Mansour, and Christophe Osswald. Adaptive strategy and decision making model for sensing-based network applications. In *2019 19th International Symposium on Communications and Information Technologies (ISCIT)*, pages 96–101. IEEE, 2019.
- [24] Alexander M Wyglinski, Maziar Nekovee, and Thomas Hou. *Cognitive radio communications and networks: principles and practice*. Academic Press, 2009.
- [25] Beibei Wang and KJ Ray Liu. Advances in cognitive radio networks: A survey. *IEEE Journal of selected topics in signal processing*, 5(1):5–23, 2010.
- [26] Spectrum Efficiency Working Group et al. Spectrum policy task force report, federal communications commission. *FCC ET Docket*, pages 02–135, 2002.
- [27] Ian F Akyildiz, Won-Yeol Lee, Mehmet C Vuran, and Shantidev Mohanty. Next generation/dynamic spectrum access/cognitive radio wireless networks: A survey. *Computer networks*, 50(13):2127–2159, 2006.
- [28] Joseph Mitola and Gerald Q Maguire. Cognitive radio: making software radios more personal. *IEEE personal communications*, 6(4):13–18, 1999.
- [29] Mohamed Abaza, Raed Mesleh, Ali Mansour, and El-Hadi M Aggoune. Cooperative mimo and multi-hop relaying techniques for free-space optical communications: a survey, 2018.
- [30] Matthew NO Sadiku, Sarhan M Musa, and Sudarshan R Nelatury. Free space optical communications: an overview. *European scientific journal*, 12(9):55–68, 2016.
- [31] Joseph Mitola. Software radio architecture: a mathematical perspective. *IEEE Journal on selected areas in communications*, 17(4):514–538, 1999.
- [32] Ian F Akyildiz, Won-Yeol Lee, Mehmet C Vuran, and Shantidev Mohanty. A survey on spectrum management in cognitive radio networks. *IEEE Communications magazine*, 46(4):40–48, 2008.
- [33] Federal Communications Commission et al. Establishment of interference temperature metric to quantify and manage interference and to expand available unlicensed operation in certain fixed mobile and satellite frequency bands. *Et Docket*, (03-237), 2003.

- [34] Cheng-Xiang Wang, Xuemin Hong, Hsiao-Hwa Chen, and John Thompson. On capacity of cognitive radio networks with average interference power constraints. *IEEE Transactions on Wireless Communications*, 8(4):1620–1625, 2009.
- [35] Chunsheng Xin, Min Song, Liangping Ma, George Hsieh, and Chien-Chung Shen. Network coding relayed dynamic spectrum access. In *Proceedings of the 2010 ACM workshop on Cognitive radio networks*, pages 31–36, 2010.
- [36] Tevfik Yucek and Huseyin Arslan. A survey of spectrum sensing algorithms for cognitive radio applications. *IEEE communications surveys & tutorials*, 11(1): 116–130, 2009.
- [37] Ju Ren, Yaoxue Zhang, Ning Zhang, Deyu Zhang, and Xuemin Shen. Dynamic channel access to improve energy efficiency in cognitive radio sensor networks. *IEEE Transactions on Wireless Communications*, 15(5):3143–3156, 2016.
- [38] Xueqing Huang, Tao Han, and Nirwan Ansari. On green-energy-powered cognitive radio networks. *IEEE Communications Surveys & Tutorials*, 17(2):827–842, 2015.
- [39] Ahmed Sultan. Sensing and transmit energy optimization for an energy harvesting cognitive radio. *IEEE wireless communications letters*, 1(5):500–503, 2012.
- [40] Linda Gaines, John Sullivan, Andrew Burnham, and Ilias Belharouak. Life-cycle analysis for lithium-ion battery production and recycling. In *Transportation Research Board 90th Annual Meeting, Washington, DC*, pages 23–27, 2011.
- [41] Sungsoo Park and Daesik Hong. Achievable throughput of energy harvesting cognitive radio networks. *IEEE Transactions on Wireless Communications*, 13(2): 1010–1022, 2014.
- [42] Jatin Singla, Rita Mahajan, and Deepak Bagai. A survey on energy harvesting cognitive radio networks. In *2018 6th Edition of International Conference on Wireless Networks & Embedded Systems (WECON)*, pages 6–10. IEEE, 2018.
- [43] Farhan Simjee and Pai H Chou. Everlast: long-life, supercapacitor-operated wireless sensor node. In *Proceedings of the 2006 international symposium on Low power electronics and design*, pages 197–202, 2006.
- [44] Aman Kansal, Jason Hsu, Sadaf Zahedi, and Mani B Srivastava. Power management in energy harvesting sensor networks. *ACM Transactions on Embedded Computing Systems (TECS)*, 6(4):32–es, 2007.



- [45] Ag Asri Ag Ibrahim, Kashif Nisar, Yeoh Keng Hsueh, and Ian Welch. Review and analyzing rfid technology tags and applications. In *2019 IEEE 13th International Conference on Application of Information and Communication Technologies (AICT)*, pages 1–4. IEEE, 2019.
- [46] Xiao Lu, Ping Wang, Dusit Niyato, and Ekram Hossain. Dynamic spectrum access in cognitive radio networks with rf energy harvesting. *IEEE Wireless Communications*, 21(3):102–110, 2014.
- [47] Yuan Gao, Haixia He, Zhixiang Deng, and Xuewu Zhang. Cognitive radio network with energy-harvesting based on primary and secondary user signals. *IEEE Access*, 6:9081–9090, 2018.
- [48] Sungsoo Park, Hyungjong Kim, and Daesik Hong. Cognitive radio networks with energy harvesting. *IEEE Transactions on Wireless communications*, 12(3):1386–1397, 2013.
- [49] Abbass Nasser, Hussein Al Haj Hassan, Jad Abou Chaaya, Ali Mansour, and Koffi-Clément Yao. Spectrum sensing for cognitive radio: Recent advances and future challenge. *Sensors*, 21(7):2408, 2021.
- [50] Wan Siti Halimatul Munirah Wan Ahmad, Nurul Asyikin Mohamed Radzi, FS Samidi, A Ismail, Fairuz Abdullah, Md Zaini Jamaludin, and MohdNasim Zakaria. 5g technology: Towards dynamic spectrum sharing using cognitive radio networks. *IEEE Access*, 8:14460–14488, 2020.
- [51] Marja Matinmikko-Blue, Seppo Yrjölä, and Petri Ahokangas. Spectrum management in the 6g era: The role of regulation and spectrum sharing. In *2020 2nd 6G Wireless Summit (6G SUMMIT)*, pages 1–5. IEEE, 2020.
- [52] Li Hong-jiang and Zhu Qi. Tradeoff between utilization and collision in cognitive radio. In *2009 International Conference on Wireless Communications & Signal Processing*, pages 1–5. IEEE, 2009.
- [53] Smit B Tripathi and Mehul B Shah. A dynamic opportunistic spectrum access mac protocol for cognitive radio networks. In *2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pages 1195–1198, 2014. doi: 10.1109/ICACCI.2014.6968388.
- [54] Sarah Eljack, M. Iqbal, and Wang Furong. A multi channel cognitive mac protocol with efficient channel reservation and collision avoidance method. *Multimedia Information Networking and Security, International Conference on*, 1:115–119, 11 2009. doi: 10.1109/MINES.2009.44.

- [55] Maxime Mroue, Abbass Nasser, Benoît Parrein, Ali Mansour, Chamseddine Zaki, and Eduardo Motta Cruz. Esco: Eligibility score-based strategy for sensors selection in cr-iot: Application to lorawan. *Internet of Things*, 13:100362, 2021.
- [56] K Seshukumar, R Saravanan, and MS Suraj. Spectrum sensing review in cognitive radio. In *2013 International Conference on Emerging Trends in VLSI, Embedded System, Nano Electronics and Telecommunication System (ICEVENT)*, pages 1–4. IEEE, 2013.
- [57] Gyanendra Prasad Joshi and Sung Won Kim. A survey on node clustering in cognitive radio wireless sensor networks. *Sensors*, 16(9):1465, 2016.
- [58] Shalli Rani and Syed Hassan Ahmed. Multi-hop routing in wireless sensor networks: an overview, taxonomy, and research challenges. 2015.
- [59] Jamal N Al-Karaki and Ahmed E Kamal. Routing techniques in wireless sensor networks: a survey. *IEEE wireless communications*, 11(6):6–28, 2004.
- [60] Saad Rizvi. Cooperative multipath routing in wireless sensor networks. In *2016 19th International Multi-Topic Conference (INMIC)*, pages 1–6. IEEE, 2016.
- [61] Mohammad Ammad Uddin, Ali Mansour, Denis Le Jeune, Mohammad Ayaz, and El-Hadi M Aggoune. Uav-assisted dynamic clustering of wireless sensor networks for crop health monitoring. *Sensors*, 18(2):555, 2018.
- [62] Pineas M Egidius, Adnan M Abu-Mahfouz, Musa Ndiaye, and Gerhard P Hancke. Data aggregation in software-defined wireless sensor networks: A review. In *2019 IEEE International Conference on Industrial Technology (ICIT)*, pages 1749–1754. IEEE, 2019.
- [63] Hairong Zhao, Wuneng Zhou, and Yan Gao. Energy efficient and cluster based routing protocol for wsn. In *2012 Eighth International Conference on Computational Intelligence and Security*, pages 107–111. IEEE, 2012.
- [64] Lan-ying Li, Xiu-li Jiang, Shenghai Zhong, and Lei Hu. Energy balancing clustering algorithm for wireless sensor network. In *2009 International Conference on Networks Security, Wireless Communications and Trusted Computing*, volume 1, pages 61–64. IEEE, 2009.
- [65] Xia Wei, Jun Lu, Yuan Zhuang, and Xianqing Ling. Adaptive clustering routing optimization for wireless sensor networks. In *2013 Ninth International Conference on Natural Computation (ICNC)*, pages 1011–1015. IEEE, 2013.

- [66] AS Raghuvanshi, S Tiwari, R Tripathi, and Nand Kishor. Gk clustering approach to determine optimal number of clusters for wireless sensor networks. In *2009 Fifth International Conference on Wireless Communication and Sensor Networks (WCSN)*, pages 1–6. IEEE, 2009.
- [67] Gino Alvarado, Carlos Bosquez, Fernando Palacios, and Luis Córdoba. Low-energy adaptive clustering hierarchy protocol and optimal number of cluster head algorithm in a randomized wireless sensor network deployment. In *2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICECCOT)*, pages 1–4. IEEE, 2017.
- [68] Ashish Rauniyar and Soo Young Shin. A novel energy-efficient clustering based cooperative spectrum sensing for cognitive radio sensor networks. *International Journal of Distributed Sensor Networks*, 11(6):198456, 2015.
- [69] Huazi Zhang, Zhaoyang Zhang, Huaiyu Dai, Rui Yin, and Xiaoming Chen. Distributed spectrum-aware clustering in cognitive radio sensor networks. In *2011 IEEE global telecommunications conference-GLOBECOM 2011*, pages 1–6. IEEE, 2011.
- [70] Ram Narayan Yadav, Rajiv Misra, and Divya Saini. Energy aware cluster based routing protocol over distributed cognitive radio sensor network. *Computer Communications*, 129:54–66, 2018.
- [71] Abdul Majid, Irfan Azam, Abdul Waheed, Muhammad Zain-ul Abidin, Taimur Hafeez, Zahoor Ali Khan, Umar Qasim, and Nadeem Javaid. An energy efficient and balanced energy consumption cluster based routing protocol for underwater wireless sensor networks. In *2016 IEEE 30th International Conference on Advanced Information Networking and Applications (AINA)*, pages 324–333. IEEE, 2016.
- [72] Ke Guan and Li-Ming He. A novel energy-efficient multi-path routing protocol for wireless sensor networks. In *2010 International Conference on Communications and Mobile Computing*, volume 3, pages 214–218. IEEE, 2010.
- [73] Burhan Gulbahar and Ozgur B Akan. Information theoretical optimization gains in energy adaptive data gathering and relaying in cognitive radio sensor networks. *IEEE Transactions on Wireless Communications*, 11(5):1788–1796, 2012.
- [74] Mingyuan Yan, Shouling Ji, and Zhipeng Cai. Time efficient data aggregation scheduling in cognitive radio networks. In *2013 IEEE Global Communications Conference (GLOBECOM)*, pages 1185–1190. IEEE, 2013.
- [75] Mingyuan Yan, Meng Han, Chunyu Ai, Zhipeng Cai, and Yingshu Li. Data aggregation scheduling in probabilistic wireless networks with cognitive radio capability.

- In *2016 IEEE global communications conference (GLOBECOM)*, pages 1–6. IEEE, 2016.
- [76] Quan Chen, Zhipeng Cai, Lianglun Cheng, and Hong Gao. Low-latency data aggregation scheduling for cognitive radio networks with non-predetermined structure. *IEEE Transactions on Mobile Computing*, 20(7):2412–2426, 2020.
- [77] Hassan Harb, Chady Abou Jaoude, and Abdallah Makhoul. An energy-efficient data prediction and processing approach for the internet of things and sensing based applications. *Peer-to-Peer Networking and Applications*, 13(3):780–795, 2020.
- [78] Ali Shokouhi Rostami, Marzieh Badkoobe, Farahnaz Mohanna, Ali Asghar Rahmani Hosseinabadi, Arun Kumar Sangaiah, et al. Survey on clustering in heterogeneous and homogeneous wireless sensor networks. *The Journal of Supercomputing*, 74(1):277–323, 2018.
- [79] Xuxun Liu. A survey on clustering routing protocols in wireless sensor networks. *sensors*, 12(8):11113–11153, 2012.
- [80] Dali Wei, Yichao Jin, Serdar Vural, Klaus Moessner, and Rahim Tafazolli. An energy-efficient clustering solution for wireless sensor networks. *IEEE transactions on wireless communications*, 10(11):3973–3983, 2011.
- [81] Jukka Lempiäinen and Matti Manninen. *Radio interface system planning for GSM/GPRS/UMTS*. Springer Science & Business Media, 2007.
- [82] Kyu-min Kang, Jae Cheol Park, Sang-in Cho, Byung Jang Jeong, Young-jin Kim, Hyoung-jin Lim, and Gi-hong Im. Deployment and coverage of cognitive radio networks in tv white space. *IEEE Communications Magazine*, 50(12):88–94, 2012. doi: 10.1109/MCOM.2012.6384456.
- [83] Aristidis Likas, Nikos Vlassis, and Jakob J Verbeek. The global k-means clustering algorithm. *Pattern recognition*, 36(2):451–461, 2003.
- [84] Kemal Akkaya and Ismail Ari. In-network data aggregation in wireless sensor networks. *Handbook of Computer Networks: LANs, MANs, WANs, the Internet, and Global, Cellular, and Wireless Networks*. Wiley Online Library, 2:1131–1146, 2007.
- [85] Dnyaneshwar Mantri, Neeli Rashmi Prasad, and Ramjee Prasad. Scheduled collision avoidance in wireless sensor network using zigbee. In *2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pages 2129–2134. IEEE, 2014.

- [86] Hassan Harb, Abdallah Makhoul, David Laiymani, and Ali Jaber. A distance-based data aggregation technique for periodic sensor networks. *ACM Transactions on Sensor Networks (TOSN)*, 13(4):1–40, 2017.
- [87] Mohamad Rida, Abdallah Makhoul, Hassan Harb, David Laiymani, and Mahmoud Barhamgi. Ek-means: A new clustering approach for datasets classification in sensor networks. *Ad Hoc Networks*, 84:158–169, 2019.
- [88] L Malathi, RK Gnanamurthy, and Krishnan Chandrasekaran. Energy efficient data collection through hybrid unequal clustering for wireless sensor networks. *Computers & Electrical Engineering*, 48:358–370, 2015.
- [89] Chittabrata Ghosh, Carlos Cordeiro, Dharma P Agrawal, and M Bhaskara Rao. Markov chain existence and hidden markov models in spectrum sensing. In *2009 IEEE International Conference on Pervasive Computing and Communications*, pages 1–6. IEEE, 2009.
- [90] Rei-Heng Cheng and Chiming Huang. The impact of the transmission power range on energy consumption for wireless sensor networks. In *2013 Fifth International Conference on Ubiquitous and Future Networks (ICUFN)*, pages 77–81. IEEE, 2013.
- [91] Abbass Nasser, Ali Mansour, Koffi-Clement Yao, Hussein Charara, and Mohamad Chaitou. Spectrum sensing for full-duplex cognitive radio systems. In *International Conference on Cognitive Radio Oriented Wireless Networks*, pages 363–374. Springer, 2016.
- [92] Remah Alshinina and Khaled Elleithy. Performance and challenges of service-oriented architecture for wireless sensor networks. *Sensors*, 17(3):536, 2017.
- [93] Ha Duyen Trung, Watit Benjapolakul, and Phan Minh Duc. Performance evaluation and comparison of different ad hoc routing protocols. *Computer communications*, 30(11-12):2478–2496, 2007.
- [94] Maryam Shafiq, Humaira Ashraf, Ata Ullah, and Shireen Tahira. Systematic literature review on energy efficient routing schemes in wsn—a survey. *Mobile Networks & Applications*, 25(3), 2020.
- [95] Charles E Perkins and Elizabeth M Royer. Ad-hoc on-demand distance vector routing. In *Proceedings WMCSA '99. Second IEEE Workshop on Mobile Computing Systems and Applications*, pages 90–100. IEEE, 1999.
- [96] Vinod Kumar Verma, Surinder Singh, and Nagendra P Pathak. Analysis of scalability for aodv routing protocol in wireless sensor networks. *Optik*, 125(2):748–750, 2014.

- [97] AAA Radwan, TM Mahmoud, and EH Houssein. Evaluation comparison of some ad hoc networks routing protocols. *Egyptian Informatics Journal*, 12(2):95–106, 2011.
- [98] Dilip Kumar Ahirwar, Prashant Verma, and Jitendra Daksh. Enhanced aodv routing protocol for wireless sensor network based on zigbee. In *International Conference on Computer Science and Information Technology*, pages 97–102. Springer, 2012.
- [99] Rajesh Sharma and Seema Sabharwal. Dynamic source routing protocol (dsr). *International Journal of Advanced Research in Computer Science and Software Engineering*, 3(7), 2013.
- [100] Chuan Xu, Zhengying Xiong, Guofeng Zhao, and Shui Yu. An energy-efficient region source routing protocol for lifetime maximization in wsn. *IEEE Access*, 7: 135277–135289, 2019.
- [101] Noman Shabbir and Syed Rizwan Hassan. Routing protocols for wireless sensor networks (wsns). *Wireless Sensor Networks-Insights and Innovations*, 2017.
- [102] Zygmunt J Haas and Marc R Pearlman. The performance of query control schemes for the zone routing protocol. *IEEE/ACM Transactions on networking*, 9(4):427–438, 2001.
- [103] Konstantinos N Amouris, Symeon Papavassiliou, and Miao Li. A position-based multi-zone routing protocol for wide area mobile ad-hoc networks. In *1999 IEEE 49th Vehicular Technology Conference (Cat. No. 99CH36363)*, volume 2, pages 1365–1369. IEEE, 1999.
- [104] Nicklas Beijar. Zone routing protocol (zrp). *Networking Laboratory, Helsinki University of Technology, Finland*, 9:1–12, 2002.
- [105] Young-Bae Ko and Nitin H Vaidya. Location-aided routing (lar) in mobile ad hoc networks. *Wireless networks*, 6(4):307–321, 2000.
- [106] Badri Nath and Dragoş Niculescu. Routing on a curve. *ACM SIGCOMM Computer Communication Review*, 33(1):155–160, 2003.
- [107] Dragoş Niculescu and Badri Nath. Trajectory based forwarding and its applications. In *Proceedings of the 9th annual international conference on Mobile computing and networking*, pages 260–272, 2003.
- [108] Yan Yu, Ramesh Govindan, and Deborah Estrin. Geographical and energy aware routing: A recursive data dissemination protocol for wireless sensor networks. 2001.

- [109] Min Chen, Taekyoung Kwon, Shiwen Mao, Yong Yuan, and Victor CM Leung. Reliable and energy-efficient routing protocol in dense wireless sensor networks. *International Journal of Sensor Networks*, 4(1-2):104–117, 2008.
- [110] Petr Musilek, Pavel Krömer, and T Bartoň. Review of nature-inspired methods for wake-up scheduling in wireless sensor networks. *Swarm and evolutionary computation*, 25:100–118, 2015.
- [111] Jamal N Al-Karaki and Ahmed E Kamal. Routing techniques in wireless sensor networks: a survey. *IEEE wireless communications*, 11(6):6–28, 2004.
- [112] Gyanendra Prasad Joshi, Seung Yeob Nam, and Sung Won Kim. Cognitive radio wireless sensor networks: applications, challenges and research trends. *Sensors*, 13(9):11196–11228, 2013.
- [113] Ghalib A Shah and Ozgur B Akan. Cognitive adaptive medium access control in cognitive radio sensor networks. *IEEE transactions on vehicular technology*, 64(2):757–767, 2014.
- [114] Mert Can Oto and Ozgur B Akan. Energy-efficient packet size optimization for cognitive radio sensor networks. *IEEE Transactions on Wireless Communications*, 11(4):1544–1553, 2012.
- [115] A Ozan Bicen, V Cagri Gungor, and Ozgur B Akan. Delay-sensitive and multimedia communication in cognitive radio sensor networks. *Ad Hoc Networks*, 10(5):816–830, 2012.
- [116] Fan Zhang, Tao Jing, Yan Huo, and Kaiwei Jiang. Outage probability minimization for energy harvesting cognitive radio sensor networks. *Sensors*, 17(2):224, 2017.
- [117] Zan Li, Boyang Liu, Jiangbo Si, and Fuhui Zhou. Optimal spectrum sensing interval in energy-harvesting cognitive radio networks. *IEEE Transactions on Cognitive Communications and Networking*, 3(2):190–200, 2017.
- [118] Subhankar Chatterjee, Santi P Maity, and Tamaghna Acharya. Energy-spectrum efficiency trade-off in energy harvesting cooperative cognitive radio networks. *IEEE Transactions on Cognitive Communications and Networking*, 5(2):295–303, 2019.
- [119] M. Almasri, A. Mansour, C. Moy, A. Assoum, D. Le Jeune, and C. Osswald. Managing single or multi-users channel allocation for the priority cognitive access. In *2020 28th European Signal Processing Conference (EUSIPCO)*, pages 1722–1726, 2021. doi: 10.23919/Eusipco47968.2020.9287628.

- [120] Zhe Chen, Zhen Hu, and Robert C Qiu. Quickest spectrum detection using hidden markov model for cognitive radio. In *MILCOM 2009-2009 IEEE Military Communications Conference*, pages 1–7. IEEE, 2009.
- [121] Thao Nguyen, Brian L Mark, and Yariv Ephraim. Spectrum sensing using a hidden bivariate markov model. *IEEE Transactions on Wireless Communications*, 12(9):4582–4591, 2013.
- [122] Xiaoshuang Xing, Tao Jing, Hongjuan Li, Yan Huo, Xiuzhen Cheng, and Taieb Znati. Optimal spectrum sensing interval in cognitive radio networks. *IEEE Transactions on parallel and Distributed Systems*, 25(9):2408–2417, 2013.
- [123] Boyang Liu, Zan Li, Jiangbo Si, and Fuhui Zhou. Optimal sensing interval in cognitive radio networks with imperfect spectrum sensing. *IET Communications*, 10(2):189–198, 2016.