LIFETIME OPTIMIZATION OF A SINGLE AND MULTIPLE SINKS WIRELESS SENSOR NETWORK ENSURING MAXIMUM COVERAGE AND CONNECTIVITY

A

Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

by

TAPAN KUMAR JAIN

(096005)

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING
JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY
WAKNAGHAT, SOLAN-173234
INDIA

MARCH 2016

Declaration

I hereby declare that the work reported in the Ph.D. thesis entitled "**Lifetime Optimization of a Single and Multiple Sinks Wireless Sensor Network Ensuring Maximum Coverage and Connectivity**" submitted at Jaypee University of Information Technology, Waknaghat India, is an authentic record of my work carried out under the supervision of Prof. Davinder Singh Saini and Prof. Sunil Bhooshan. I have not submitted this work elsewhere for any other degree or diploma.

Tapan Kumar Jain

Department of Electronics and Communication Engineering

Jaypee University of Information Technology

Waknaghat, Solan-173234

India

Certificate

This is to certify that the thesis entitled, "**Lifetime Optimization of a Single and Multiple Sink Wireless Sensor Network Ensuring Maximum Coverage and Connectivity**" which is being submitted by **Tapan Kumar Jain** in fulfillment for the award of degree of **Doctor of Philosophy** in Electronics and Communication Engineering by the Jaypee University of Information Technology, is the record of candidate's own work carried out by him under my supervision. This work has not been submitted partially or wholly to any other University or Institute for the award of this or any other degree or diploma.

Prof. Davinder Singh Saini
Professor and Head
Department of Electronics and Communication Engineering
Chandigarh College of Engineering and Technology
Chandigarh -160019 India

Prof. Sunil Bhooshan
Professor and Head
Department of Electronics and Communication Engineering
Jaypee University of Information Technology
Waknaghat, Solan-173234 India

Acknowledgment

First and the foremost I would like to thank my supervisors, Dr. Davinder Singh Saini and Prof. Sunil Bhooshan, for their guidance during my research and doctoral study at Jaypee University of Information Technology (JUIT), Waknaghat. Their perpetual energy and enthusiasm in research has motivated and inspired me to discipline myself in efficiently executing my multiple responsibilities simultaneously. In addition, they were always accessible and willing to help me with my queries and doubts during my research, resulting in its rewarding completion along with my accomplishment.

I was always delighted to interact with Prof. S. C. Saxena, Prof. Yaj Medury, Brig. Balbir Singh (Retd), Prof. Samir Dev Gupta, Prof. T. S. Lamba, and Prof. G. Singh. Their insights and suggestions were instrumental in adding several interesting dimensions to my study. I would like to acknowledge the members of my thesis committee, Prof. Harinder Singh, Dr. Shruti Jain, and Dr. Neeru Sharma for their support and suggestions which improved the quality of thesis. The entire department of Electronics and Communication Engineering at JUIT deserve my special gratitude for broadening my perspective on the practical aspects of my chosen research subject.

My deepest gratitude goes to my wife, Pooja and my son, Vardhan for their unflagging love and support throughout my research; this dissertation is simply impossible without them. Unconditional love and support from my parents and all family members boosted my morale as and when I felt lacking. Special thanks to my father in law, who was always there whenever I needed support and motivation.

I owe my gratitude towards Acharya Pushpadant Sagar and Prasann Sagar ji Maharaj who inspired me to be a responsible individual and illustrated the power of knowledge, humility and forgiveness.

Contents

List of Figures

List of Tables

List of Algorithms

Glossary

ACQUIRE	Active Query Forwarding in Sensor Networks
AODV	Ad-hoc on-demand Vector Routing
APTEEN	Adaptive Periodic Threshold Sensitive Energy Efficient Sensor Network Protocol
BS	Base Station
BVGF	Bounded Voronoi Greedy Forwarding
CH	Cluster Head
CHR	Cluster-Head Relay Routing
CHS_msg	Cluster Head Selection message
CLP	Concentrator Location Problem
DA	Data aggregation
DCT	Discrete Cosine Transform
DSR	Dynamic Source Routing
DTE	Direct Transmission Energy
EAD	Energy-Aware Data-Centric Routing
FND	First Node Died Out
GAF	Geographic Adaptive Fidelity
GEAR	Geographic and Energy-Aware Routing
GeRaF	Geographic Random Forwarding
GPS	Global Positioning System
HAC	Hierarchical agglomerative clustering
IDSQ	Information-Driven Sensor Query
LEACH	Low Energy Adaptive Clustering Hierarchy
LND	Last Node Died Out
MECN	Minimum Energy Communication Network
MEMS	Micro Electro Mechanical Systems
MH-MS	Multiple Hops Multiple Sink Routing
MH-SS	Multiple Hops Single Sink Routing
MTE	Minimum Transmission Energy
NL	Network lifetime

PEGASIS	Power-Efficient Gathering in Sensor Information Systems
QoS	Quality of Service
RSS	Received Signal Strength
RSSI	RSS Indicator
SEAD	Scalable Energy-Efficient Asynchronous Dissemination
SH-MS	Single hop multiple sink routing
SH-SS	Single hop single sink routing
SMECH	Small Minimum-Energy Communication Network
SNR	Signal to Noise Ratio
SPIN	Sensor Protocols for Information via Negotiation
TEEN	Threshold Sensitive Energy Efficient Sensor Network Protocol
WEMS	Wireless Environment Monitoring System
WSN	Wireless Sensor Network

Abstract

Wireless sensor network (WSN) consists of a large number of small sensing nodes, capable of sensing any physical quantity like temperature, humidity, rainfall etc. WSNs have created new opportunities across the spectrum of human endeavors including engineering design and manufacturing, monitoring and control of systems. Involvement of restrained resources in the deployment of WSNs makes it a subject of concern. So its usage needs to be energy efficient in order to maximize operational life of network.

Energy efficient WSN is used to monitor the environment and warns for a mishappening occurring in future. A merging algorithm is proposed which minimizes the global energy usage by making sensing nodes to work in sleep and active mode. Merging outperforms the conventional network's energy usage by uniformly distributing the load of sensing in the nodes falling in the same sensing range. Another approach is to have efficient clustering without requiring the global knowledge of network by using the bottom-up called the Hierarchical agglomerative clustering (HAC) which consequently reduces the energy consumption to a great extent.

Maximum coverage and connectivity is achieved by employing the concept of unique nodes. In the proposed work, there is no possibility of having any outlier, as all the unique nodes are connected to some or the other cluster head (CH). In WSNs, the communication cost is often several orders of magnitude larger than the computation cost, thus the CHs perform data aggregation to reduce the amount of data to be transmitted.

Multiple sink WSN has an edge over the single sink where very less energy is utilized in sending the data to the sink, as the number of hops are reduced.The network lifetime optimization is achieved by re-structuring the network by modifying the neighbor nodes of a sink.

1. Introduction

A wireless sensor network (WSN) consists of small low energy sensing nodes capable of sensing a phenomenon and sending the data to the sink. Advances in the field of wireless communication, Micro Electro Mechanical Systems (MEMS) technology have led to the development of low cost ,mulifunctional [1] tiny sensor nodes which consume less power. WSNs are basically data gathering networks in which data are highly correlated and the end user needs a high level description of the environment sensed by the nodes[1]. WSNs are deployed to monitor physical events or the state of physical objects such as bridges in order to support appropriate reaction to avoid potential damages [2]. The nodes and the related protocols in a WSN should be designed to be extremely energy efficient as battery recharging may be impossible [3].

A sensor node cannot function without the power unit as this unit supports all the other functions in a sensor node. With every sensor node there are two very crucial associated parameters namely:

• Sensing range (R_s): It is the maximum distance upto which sensing node can sense a phenomenon.

• Transmission range (R_t): It is the optimum distance over which a sensor node can transmit data.

The basic purpose of WSN is to sense a phenomenon and send it back to the Base Station (BS). No human intervention takes WSN to the whole new level and because of this the WSNs have various applications and can reach where humans can't, like in battle field surveillance, disaster prone areas, detection on a gas leak in nuclear power plants etc. The sensor nodes are deployed in a specific region (inside the phenomenon or close to it) and are deployed randomly or manually. Typically in WSN, nodes coordinate locally to gather and process data and send it to common sink[4]. The BS is assumed to contain infinite energy. Nodes in WSN are driven by limited battery power and are not rechargeable. The energy contained in a sensor node is consumed in various processes such as sensing, processing and communication. Since, the nodes have limited operating power, energy optimization becomes extremely imperative for a WSN to function for a longer time. Energy of nodes and energy usage are key factors in determining the lifetime of whole network.

One of the most important challenges of a WSN design is to develop a protocol so that the randomly deployed numerous sensing units behave in a collaborative, cooperative, coordinated and organized way. Each node needs to send the data to its neighbour nodes and finally to the sink. Network routing protocol design becomes quite critical in the case of WSNs as compared to conventional communication networks [5]. Among various proposed network routing protocols, hierarchical routing protocols or clustering greatly contribute to system scalability, lifetime, and energy efficiency [6].

Many researches investigating the field of energy optimization of WSN and many energy efficient routing protocols such as the minimum transmission energy and LEACH[3] (low energy adaptive clustering hierarchy) protocol have been proposed, where the concept of clustering, cluster head (CH) and techniques like data aggregation and data fusion are used which help in increasing the efficiency of the network[7].

Clustering is a useful mechanism in WSNs which helps to cope with scalability problems. When combined with data aggregation, clustering may increase the energy efficiency of the network. Moreover, by assigning a special role to the CHs, clustering makes the network all the more robust and vulnerable to attacks [8]. In direct communication WSN, the sensor nodes directly transmit their sensed data to the sink without any coordination between the two. However, in cluster based WSNs, the network is divided into clusters. Each node exchanges its information only with its CH which transmits the aggregated information to BS[9]. Data aggregation at CHs causes a significant reduction in the amount of data sent to the BS and results in saving both energy and bandwidth resources [10]. Effective clustering algorithm leads WSN to operate efficiently. However, the major challenges are the equal distribution of each cluster over the entire sensor network and the energy dissipation caused by the frequent information exchange between selected CH and nodes in the cluster in every setup phase of cluster formation[11].

The network lifetime in a WSN can be defined as the time at which the first node runs out of its energy. A lot of energy of the WSN is utilized in performing various actions

1. Clustering the network,
2. Cluster head selection
3. Finding the optimum routing path,
4. Data transmission from one node to another using single or multiple hops,
5. Data aggregation and interpretation.

The network lifetime depends on all the above factors. Less the energy utilization, and more the network lifetime.

The most important phase of cluster-based routing protocols is the CH selection procedure that ensures uniform distribution of energy among the sensors, and consequently increasing the lifespan of a sensor network [12, 13]. Once the CHs are identified, they form a backbone network to periodically collect, aggregate, and forward data to the BS using the minimum energy (cost) routing. This method significantly enhance the network lifetime compared to other known methods. The major challenges include equal distribution of each cluster over the entire sensor network and the energy dissipation caused by the frequent information exchange between selected CH and nodes in the cluster in every setup phase of cluster formation [11, 4]. If CH is selected on the basis of the concept of maximum number of nodes connected, then it may happen that one or more unique nodes are not connected to any of the selected CHs. In such case some outliers may be created which are not connected to any of the CHs, although they are in the transmission range. This algorithm deals with the CH selection based on the unique node concept. A unique node is the one which is not connected to any other CHs. The algorithm proposed in this research work uses two parameters, namely, number of neighboring nodes and the residual energy for the selection of CH in WSN.

Most of the current clustering protocols are top-down approaches, which first formulate a global knowledge of a WSN, specifying but not detailing the first-level nodes. Based on this global knowledge, the protocols first build the upper level of clusters by selecting certain nodes as CHs. Then they group the rest of the nodes into the designated cluster as cluster members. Many algorithms randomly select the CHs. In such a case, it might happen that the CH may have lower energy than its member nodes. Such CH may died out quickly which usually results in low quality of the clusters.

The motivation of our research is to provide efficient clustering without requiring the global knowledge of network by reversing the clustering approach from top-down to bottom-up[14]. With the bottom-up approach, sensing nodes build clusters before they select CHs. In this manner, the bottom-up approach can be a better way to implement self-organization, scalability and flexibility. Such a bottom up clustering is called agglomerative clustering.

Usually location data is used to calculate the distance between the sensing nodes to perform clustering. This type of clustering is called quantitative agglomerative hierarchical clustering. But the location data may not always be available[15] due to reasons like Global Positioning System (GPS) failures or cost involved or time taken to calculate the exact location of the sensors. To avoid this, Received Signal Strength (RSS) or RSS Indicator (RSSI) is used to find out the distance between the nodes. This type of clustering is called qualitative agglomerative hierarchical clustering. The current research work compares the different agglomerative protocols for qualitative and quantitative data.

Maximum energy consumption takes place in communicating the data from the nodes to the sink [16]. To minimize the energy consumption while sending the data to the sink, multiple sinks are used. The proposed work intends to reduce the energy utilization by deploying multiple sinks. Generally a WSN is based on many-to-one communication concept of having many sensors transferring the data to a common single sink. The proposed work demonstrates the importance of multiple sink in a WSN [17]. The most important advantage of having multiple sinks is to shorten the routing path between a sensing node and the sink. If the area to be sensed in very huge and the sensors are randomly deployed, then it happens that even a very effective routing protocol fails. Even if clustering is used, the nodes waste a lot of energy in sending the data to long distances. If the sink or the next CH is within the transmission range of the sensor then the energy consumed in free space is related with distance as [18](Section 4.2)

$$E \propto d^2 \tag{1.1}$$

where d is the distance from source to destination node. But if the sink or the next CH is not within the transmission range of the sensor node then the link follows multipath propagation model for wireless scenario. In such a case, energy consumed is related with distance as [18](Section 4.2)

$$E \propto d^4 \tag{1.2}$$

So, to save energy, it becomes very important to reduce the distance between the node and the sink or the next CH. The above problem is solved by using multiple sinks in a WSN. The current research work concentrates on two basic aspects related to multiple sinks WSN, namely:

1. Finding optimal number of sinks
2. Finding the position of sinks

Multiple sinks deployment makes the WSN robust so that even if one of the sinks fail, other sinks can take the charge and prevent the WSN from failing[19].

As there are multiple sinks, the distance from the node to the sink reduces, thus there is no need of multiple hops. Multiple sinks reduce the distance the sensed data needs to travel and hence correspondingly reduce the energy consumption considerably. Another disadvantage of a single sink WSN is that of energy imbalance between the nodes close to the sink and the ones which are far off [20]. The network is restructured by modifying the number of nodes connected to a sink[21]. The current research work proposes an algorithm for network restructuring in a multiple sink WSN so as to reduce the energy consumption and increase the network lifetime. This energy balancing through network re-structuring optimizes the network lifetime. The number of non-connected nodes is considerably reduced. The implementation is done in MATLAB which demostrates the aforesaid statements.

1.1. Motivation and Objective

Wireless sensor networks (WSNs) are being used in a variety of application like border surveillance, ocean bed tracking, military field etc. In all such applications, sensors are deployed randomly and are unattended for a very long time. The network should be able to work for a long duration without any human intervention. The nodes should be able to cover maximum possible area. The motivation of the current research work concentrates on such a WSN which are energy efficient and have maximum coverage. Our objective is to design WSN to ensure

- Minimal energy consumption,
- Maximum coverage and connectivity
- Optimized network lifetime.

1.2. Contribution of Thesis

The thesis contributes by proposing novel algorithms for single and multiple sinks WSN. The algorithms aim to minimize the energy consumption, maximize the coverage and connectivity and optimize the network lifetime. The thesis is divided into two parts. The first part contains the work proposed for single sink and the second part contains the algorithms proposed for multiple sinks.

Single Sink

A Wireless Environmental Monitoring System (WEMS) is proposed using data aggregation in a bidirectional hybrid protocol [4]. A framework for energy efficient routing protocol is proposed for homogeneous WSN using sensing range as the parameter [9]. If two nodes are sensing the same area or if the node is falling in the sensing range (R_s) of another node connected to the same CH then it is of no use to sense same data using multiple sensors. This reduces the energy efficiency of the network and as we have limited available operating power for wireless networks, this will be a major reason for shortening the lifetime of network. An improvement over this approach is the proposed merging algorithm. Performance analysis of hierarchical agglomerative clustering is performed in a WSN using quantitative and qualitative data [14]. A cluster head selection algorithm is proposed for a homogeneous WSN ensuring full connectivity with minimum number of isolated nodes [13].

Multiple Sink

The most important advantage of having multiple sinks is to shorten the routing path between a sensing node and the sink. Two algorithms have been proposed for multiple sinks. One is for increasing the lifetime of a WSN using Multiple Sinks [19]. The second is for lifetime optimization through energy balancing [21]. The network is restructured by modifying the number of nodes connected to a sink.

1.3. Organization of Thesis

The second chapter of the thesis deals with the review of literature. It discusses the various terms, methodologies and protocols proposed by different researchers in the field of WSN. It also talks about the coverage and connectivity issues and their remedies. The third chapter deals with a novel clustering algorithm for Wireless Environmental Monitoring System. It is an application based algorithm having features which are suitable for a system used to monitor critical environmental conditions, such as temperature, humidity, intrusion, and smoke. Another clustering algorithm is proposed which uses sensing range as a parameter for clustering. The proposed algorithm solves the major constraint of WSN i.e. energy consumption. The algorithm makes the WSN a quite energy efficient network as the concept of merging in a multi-hop network is proposed in which if two sensing nodes are falling in the same sensing range then they will work in sleep and active mode as a result of which less energy will be used.

The author proposes a Hierarchical Agglomerative Clustering algorithm in the fourth chapter. The motivation of the research is to provide efficient clustering without requiring the global knowledge of network by using the bottom-up called the Hierarchical agglomerative clustering (HAC). If the distance between the sensing nodes is calculated using their location then it's quantitative HAC. If the received signal strength is used to calculate the distance between the nodes then it's qualitative HAC.The comparison of the various agglomerative clustering techniques applied in a WSN is also done. The author proposes a novel approach to ensure full Coverage and Connectivity in Single Sink in the fifth chapter. The second part of thesis deals with multiple sinks. The chapters 6 and 7 deal with algorithms to increase life time in multiple sinks. As the energy consumption of the network is reduced, the lifetime of the network is considerably increased. Every chapter's introduction contains the critical literature review and the motivation to reach to the novel approach.

2. Theoretical Background

2.1. Introduction

This chapter describes the work done in the field of Wireless Sensor Network (WSN) and also specifies why this topic was chosen by the author for his research work. Section 2.1 deals with clustering. The section 2.2 and 2.3 deals with Routing and the different routing protocols proposed by different researchers respectively. The section 2.4 discusses the design parameters of WSN. The next section deals with the coverage issues and the ways to handle these issues in WSN. The section 2.6 deals with connectivity issues and their solutions in WSN. The subsequent section introduces the concept of having multiple sinks. With multiple sinks, the network becomes many to many in contrast with traditional many to one scenario. The section 2.8 introduces the radio model. Finally, the last section summarizes the work proposed in different chapters of the thesis based on the critical literature review mentioned in sections 2.2 to 2.8.

2.2. Routing in WSN

WSN normally consists of a large number of sensing nodes deployed in a region to sense any physical quantity like temperature, pressure, humidity etc[1]. Such sensing nodes are capable of sensing, processing and transmitting the data to a base station or a sink. If the sink is within the transmission range of the node, it transfers the data in single hop. But if the sink is far off, the node sends the data in multiple hops. Sensor nodes have a limited transmission range as well as limited processing and storage capabilities. When the nodes have to send the data in multiple hops, a routing path must be established between the node and the sink. The routing path is determined by a routing protocol. The routing protocols have to ensure reliable multi-hop communication under the limited resources of a WSN. The design space for routing algorithms for WSNs is quite large and we can classify the routing algorithms [22] for WSNs in many different ways. Different routing protocols are depicted in the Figure 2.1. They are discussed in detail in the subsequent section.

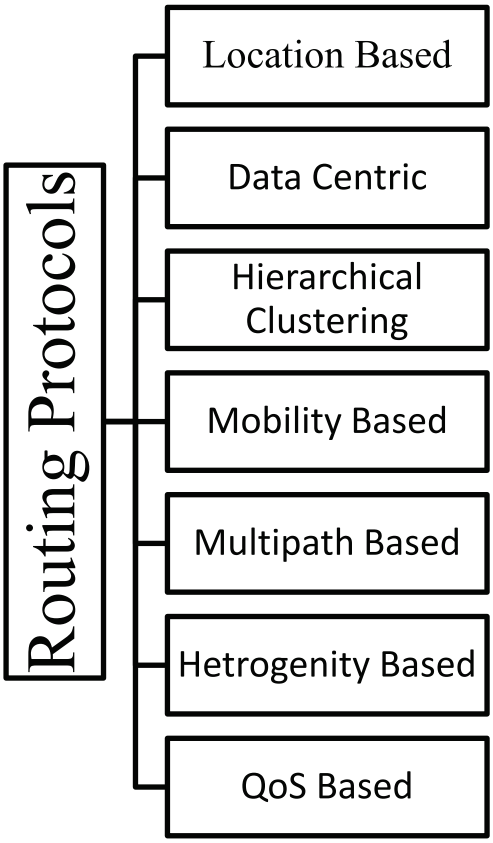

Figure 2.1.: Routing protocols

Table 2.1.: Routing Protocols

Type	Name of the Protocols
Location Based	GAF[29], GEAR[30], Span [31], TBF[32] , BVGF[33], GeRaF[34], (MECN)[35], (SMECN)[36]
Data Centric	SPIN[38, 39], Directed diffusion [40]
Hierarchical Clustering	Rumor routing[41], Cougar protocol[42] ACQUIRE[43], EAD[44] LEACH[3], Leach[47], Leach-M[48], Leach-A[49], Leach-B[50], Leach-C[51], Leach-E[52], Leach-F[53] Leach-L[54], PEGASIS [55], TEEN[57], APTEEN[58]
Mobility Based	Joint Mobility and Routing Protocol[59], Data MULES Based Protocol[60], SEAD[61]
Multipath Based	Sensor-disjoint multipath routing[62], Braided multipath[63], N-to-1 Multipath Discovery [64]
Hetrogenity Based	IDSQ[64], CHR[65]
QoS Based	SAR[66], SPEED[67], DSR[68], AODV[69]

2.3. Routing Protocols

Many routing algorithms were developed for wireless networks in general. Many researchers have written survey papers comparing the different routing protocols [23, 24, 25, 26, 27]. All major routing protocols proposed for WSNs may be divided into seven categories [28] as depicted in Table 2.1 They will be described in the subsequent sub sections.

2.3.1. Location Based Protocols

In location based protocols, the sensing nodes are addressed with the help of their locations. Such protocols are equipped with a hardware to determine the location of the node like a GPS unit or a localization system[37]. Geographic Adaptive Fidelity (GAF): GAF[29] is an energy-aware routing protocol primarily proposed for MANETs, but can also be used for WSNs because it favors energy conservation. Geographic and Energy-Aware Routing (GEAR) [30] is an energy-efficient routing protocol proposed for routing queries to target regions in a sensor field. Span[31] is a routing protocol which is motivated by the fact that the wireless network interface of a device is often the largest consumer of power. Hence, it's better to turn the radio off during idle time.TBF[32] is a routing protocol which on the basis of the location information of its neighbors, has a forwarding sensor which makes a greedy decision to determine the next hop that is closest to the trajectory fixed by the source sensor.

Bounded Voronoi Greedy Forwarding (BVGF) [33]uses the concept of Voronoi diagram [6]. The sensors eligible for being the next hops are the ones whose Voronoi regions are traversed by the segment line joining the source and the destination. Geographic Random Forwarding (GeRaF) was proposed by Zorzi and Rao [34]. In this protocol, there is no guarantee that a sender will always be able to forward the message toward the sink. This is the reason why GeRaF is also called best-effort forwarding.

Minimum Energy Communication Network (MECN)[35] is a self-reconfiguring protocol which maintains the network connectivity, even though the sensors are mobile. Small Minimum-Energy Communication Network (SMECN)[36]is an improvement over MECN. The sensors discover its immediate neighbors by broadcasting a message using some initial power that is updated incrementally. Such protocols have costly sensing nodes due to the installed GPS unit.

2.3.2. Data Centric Protocols

In data centric protocols, the sensors send their data to the sink in multiple hops. The intermediate sensors perform some form of data aggregation on the data coming from multiple sending nodes and send the aggregated data to the sink. A lot of energy is wasted in transmission of data. Thus the process of data aggregation results in energy savings.Sensor Protocols for Information via Negotiation (SPIN)[38, 39] solves the problem of implosion and overlap. The SPIN protocols are based on two key mechanisms namely negotiation and resource adaptation. The sensors negotiate with each other before any data transmission to avoid injecting non-useful and redundant information in the network. SPIN uses meta-data as the descriptors of the data that the sensors want to transmit. The meta-data avoids the occurrence of overlap as the sensors can name the interesting portion of the data they want to get.

Directed diffusion [40]is used for sensor query dissemination and processing. The main elements of this protocol are data naming, interests and gradients, data propagation, and reinforcement. Rumor Routing: Rumor routing is a logical compromise between query flooding and event flooding app schemes [41]. The Cougar protocol[42] has a query layer which allows the user to query the data from the WSN. The user need not know which of the sensors has processed its query.The protocol also uses in-network processing which in turn helps in reducing the energy consumption.

Active Query Forwarding in Sensor Networks (ACQUIRE)[43] provides query optimization which enables the sensors to answer complex queries through simple several answers. Each sub query is answered through a database stored in the relevant sensors. Energy-Aware Data-Centric Routing (EAD)[44] provides high level of in-network processing and traffic relaying.

2.3.3. Hierarchical Protocols

Considering the constraints of a WSN, energy-aware routing and data gathering protocols should be developed [45]. Grouping of sensors in clusters offers the said objectives. The

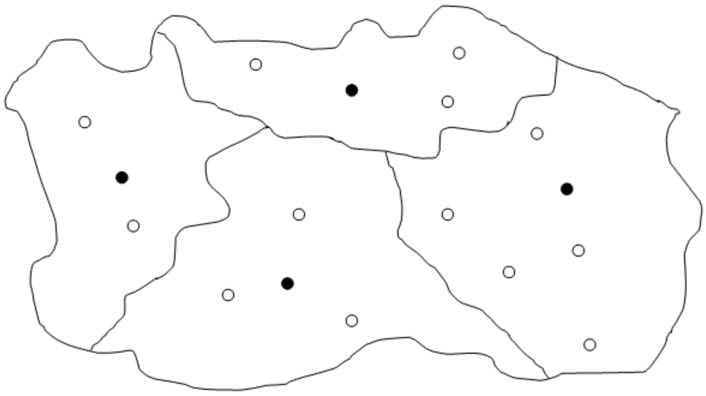

Figure 2.2.: Cluster creation in WSN

Table 2.2.: Comparison of different LEACH protocols I

Routing Protocol	Mobility	Scalability	Self Organisation	Distributed	Hop Count
LEACH	Fixed BS	Limited	Yes	Yes	Single
LEACH-S	Fixed BS	Good	Yes	Yes	Single
Multi-hop LEACH	Fixed BS	Very Good	Yes	Yes	Multi
LEACH-M	Mobile BS	Very Good	Yes	Yes	Single
LEACH-A	Fixed BS	Good	Yes	Yes	Single
LEACH-C	Fixed BS	Good	Yes	No	Single
LEACH-B	Fixed BS	Good	Yes	Yes	Single
LEACH-F	Fixed BS	Limited	No	No	Single
LEACH-L	Fixed BS	Very Good	Yes	Yes	Multi
LEACH-E	Fixed BS	Very Good	Yes	Yes	Single

clusters have a Cluster Head (CH) which performs special task of data aggregation and fusion. The research in clustering in WSN deals with the identification of CHs or finding the optimal routing path ensuring minimal energy consumption. A simple illustration of cluster formation in a region to be sensed is depicted in the Figure 2.2.

Some of the clustering protocols are explained here in brief. Low-energy adaptive clustering hierarchy (LEACH) is the first and most popular energy-efficient hierarchical clustering algorithm for WSNs [3]. In this protocol the CHs send the data directly to the sink in single hop. The CHs perform data aggregation. The CH is selected randomly among the nodes in the cluster. The comparison of different Leach protocols Leach-S[46], Multi hop Leach[47], Leach-M[48], Leach-A[49], Leach-B[50], Leach-C[51], Leach-E[52], Leach-F[53] and Leach-L[54] is shown in the Tables 2.2 and 2.3.

All the mentioned LEACH protocol variations are hierarchical routing protocols with randomized cluster head selection and they use data aggregation techniques. These protocols were designed by different researchers having one or the other benefit depending upon the

Table 2.3.: Comparison of different LEACH protocols II

Routing Protocol	Centralised	Energy Efficient	Homogeneous	Use of location information
LEACH	No	High	Yes	No
LEACH-S	Yes	V. High	Yes	No
Multi-hop LEACH	No	V. High	Yes	Yes
LEACH-M	No	V. High	Yes	Yes
LEACH-A	No	V. High	No	No
LEACH-C	Yes	V. High	Yes	Yes
LEACH-B	No	V. High	Yes	Yes
LEACH-F	Yes	V. High	Yes	Yes
LEACH-L	No	V. High	Yes	Yes
LEACH-E	No	V. High	No	Yes

application for which it was used. In one type of LEACH one parameter was better than the other LEACH variation and in other protocol some other parameter was made better. There is always a trade off between the different design parameters of WSN which solely depends upon the application for which the network is going to be used.

Power-Efficient Gathering in Sensor Information Systems (PEGASIS) [55]forms chains of sensors and only one node in that chain transmits the data to the sink. The sensors transmit the data to its neighbors instead of sending it to the CH. Instead of selecting the CH at random as in Leach, Hybrid, Energy-Efficient Distributed Clustering (HEED) [56]protocol selects a CH on the basis of residual energy.

Threshold Sensitive Energy Efficient Sensor Network Protocol (TEEN)[57]. The closer nodes form clusters and this process goes on the second level until the sink is reached. TEEN is not suitable for applications where periodic reports are required since the user may not get any data at all if the thresholds are not reached. Adaptive Periodic Threshold Sensitive Energy Efficient Sensor Network Protocol (APTEEN)[58] is an advanced TEEN protocol which senses critical as well as periodic data. It removes the drawbacks of TEEN. Hierarchical clustering is explained in detail later.

2.3.4. Mobility Based Protocols

In such protocols, the sinks are mobile. The routing protocols should be able to transmit the data to the mobile sink. In the case of static sinks, the CHs or the nodes which are close to the sink and highly loaded and get energy depleted faster as compared to other sensing nodes. To avoid this problem, the sinks are made mobile so that the nodes close to the sink are always changing and thus the chances of nodes getting die out decreases. As a result the overall network lifetime increases. Joint Mobility and Routing Protocol[59] considers mobile trajectory to be concentric circles or annuli. Data MULES Based Protocol proposes to address the need of guaranteeing cost-effective connectivity in a sparse network [60]. Scalable Energy-Efficient Asynchronous Dissemination (SEAD) is a self-organizing protocol consists of three main components namely dissemination tree (d-tree) construction, data dissemination, and maintaining linkages to mobile sinks. Dynamic Proxy Tree-Based Data Dissemination [61]form a tree between the source sensor and the multiple sinks which wish to receive the data. Each source is represented by a stationary source proxy and each sink is represented by a stationary sink proxy. But, in such protocols , extra energy is consumed in finding out the location of the nodes and the sink every time the sensing node needs to send the data.

2.3.5. Multipath Based Protocols

The routing of data from the sensing nodes to the sink is single path or multi path. In the case of single path, the node finds out the shortest path to the sink and sends the data through that path. In the case of multi path, the nodes finds 'k' shortest paths to the sink and then the node distributes its load evenly on those multiple paths. Thus the data reaches the sink on many paths as a result of which the energy consumption reduces whereby increasing the network lifetime. Sensor-disjoint multipath routing[62] protocol helps in finding a small number of alternate paths that have no sensor in common with each other or with the primary path. The primary path is the best route where as the alternate paths are less desirable because of their longer latency. Braided multipath[63] first forms the first primary path. Then for every node on the primary path, the alternate paths are constructed from those nodes to the sink. These paths are called idealized braided multipaths. N-to-1 Multipath Discovery [64] is based on the simple flooding originated from the sink. It is composed of two phases, branch aware flooding and multipath extension of flooding.

2.3.6. Heterogeneity Based Protocols

In heterogeneous WSNs, there are two types of sensors namely line-powered sensors and battery powered sensors. The line powered ones have no energy constraint, and the second ones have limited lifetime. The routing protocols should be chosen such that their available energy is efficiently managed among the two types of nodes. Information-Driven Sensor Query (IDSQ)[64]has a subset of sensors active when there are interesting events to report to the sink in some parts of the network. Cluster-Head Relay Routing (CHR)[65] uses two types of sensors to form a heterogeneous network. It has a large number of low-end sensors, denoted by L-sensors, and a small number of powerful high-end sensors, denoted by H-sensors. Both the H and L sensors are static and know their locations.

Table 2.4.: Design Factors for Protocol

Design Factor	Property
Data delivery model	Hybrid
Data aggregation	Possible
Power usage	Low
Mobility	Possible
Scalability	Good
Security	Possible
Topology	Self organizing and random

2.3.7. QoS Based Protocols

Apart from minimization of energy consumption, it is also important to consider the requirements for quality of service (QoS) with respect to reliability, delay and fault tolerance in WSNs. These protocols find a balance between energy consumption and the QoS requirements. Sequential Assignment Routing (SAR)[66]is a table-driven multi-path approach which tries to achieve energy efficiency and fault tolerance. The routing decision is dependent on factors like QoS on each path, energy resources and priority level of packet. SPEED [67] is another QoS routing protocol which provides soft realtime end-to-end guarantees. In this protocol, each node maintains its neighbor information and uses geographic forwarding to find the routing paths. SPEED ensures a certain speed for each packet to be sent to the sink. When compared to Dynamic Source Routing (DSR) [68] and Ad-hoc on-demand vector routing (AODV)[69], SPEED performs better with respect to end-to-end delay and miss ratio. In Energy-Aware QoS Routing Protocol , realtime traffic is generated by imaging sensors [70]. It finds a least cost and energy efficient path that meets certain end-to-end delay.

2.4. Design Parameters of a WSN

After comparing the existing routing protocols with the consideration of several design factors like Scalability, Power Usage, Mobility, Over-heads, Query-based, Data Aggregation and Localization, the parameters to be included in designing a new routing protocol are listed in Table 2.4. Ideally the new proposed routing protocol should follow the parameters listed in the Table 2.4, but due to the inter dependency of parameters, it is not possible to achieve all the ideal characteristics in a routing protocol.

Table 2.5.: Coverage Issues

Issues	Description
Coverage types	Depending upon the application either there is a need to monitor the entire region called full coverage , or track only a certain area of the network or sense only some set of targets.[71-75]
Deployment	The sensors are either deployed at pre-decided position in the network i.e. deterministic, or they are randomlydeployed. [76-78]
Node types	The node types can be homogeneous or heterogeneous [79 and 80]
Constraints	The constraints in designing the WSN is employed in either Energy consumption of the WSN, lifetime of the network etc.[81]
Dimensional coverage	WSN can be two dimensional or three dimensional space. [82]

2.5. Coverage in WSN

Coverage in WSN is defined as how much and for how long the sensors are able to sense the physical phenomenon in the region they are deployed to be sensed. The issues and the approaches for solving the problem of coverage is discussed in the subsequent sub sections.

2.5.1. Issues with Coverage in WSN

There are several issues while considering the coverage in WSN. Some of them are listed in Table 2.5:

Coverage types

Depending upon the application either there is a need to monitor the entire region called full coverage [71], or track only a certain area of the network or sense only some set of targets. Full or blanket coverage means that each and every point of the area is covered by atleast one sensor. The research paper[72] focuses on the problem of finding the minimum number of sensors that maintain full coverage. Given a region R containing sensors, if each crossing point in R is covered by at least one other sensor in R, then R is said to be completely covered. A crossing point is an intersection point of the two sensing disks of two neighboring sensors, or that of a sensing disk and the boundary of region R. The paper [73] takes one step further and proves that if all the crossing points in the region R are k-covered, then R is k-covered. It also proves that if all the crossing points inside the sensing range are at least k-covered, then a node can be in the sleep mode or inactive. Thus there are two coverage problems:-

1. Given the sensing range R of sensors, how to place the sensors so that the the number of sensors N needed for full coverage minimized

2. Given the number of available sensors N, how to place the sensors so that the sensing range R needed for full coverage is minimized

The current research work concentrates on clustering so as to have maximum coverage. Target coverage observes a fixed number of targets[74, 75].

Deployment

The sensors are either deployed at pre-decided position in the network i.e. deterministic [76], or they are randomly deployed[77] . The first case is generally used with costly motes and the major problem is that of placement, determining the exact locations of the sensing nodes so that the coverage is maximum. The second case covers inexpensive small motes and the most important problem in this case is that of density control. The deployment of sensors in WSN can be dense or sparse. The dense deployment is usage of large number of sensors and is used in critical application when all the events need to be tracked. Sparse deployment is the organization of less sensors. After deployment the sensors may be static or mobile. All these factors depend upon the application for which the WSN is employed. In the paper[78], the lifetime of a WSN is defined as the period starting from the deployment time until the WSN fails to satisfy its requirements (including coverage, connectivity and success transmission rate).The author produces a deployment model which affects the coverage. The current research work deals with random deployment and the nodes are static after placement.

Node types

The sensor nodes selected for the WSN are homogeneous or heterogeneous as shown in the Figure 2.3. Homogeneous nodes WSN have all the nodes of the same type with similar properties like sensing range, transmission range and energy. M.Gupta deals with the coverage of a homogeneous WSN[79]. Heterogeneous nodes WSN have nodes of different types with different properties. Guan Zhi-yan researches on coverage for such WSN[80].

Constraints

The most important constraint in designing a WSN is that of energy. The nodes are small with limited source of energy. The lifetime of a network is directly proportional to the residual energy of the node. Cardei and Wu, in their research paper have presented a summary of different approaches to energy efficient coverage problems[81]. The author has proposed algorithms to keep the energy consumption of the network as low as possible.

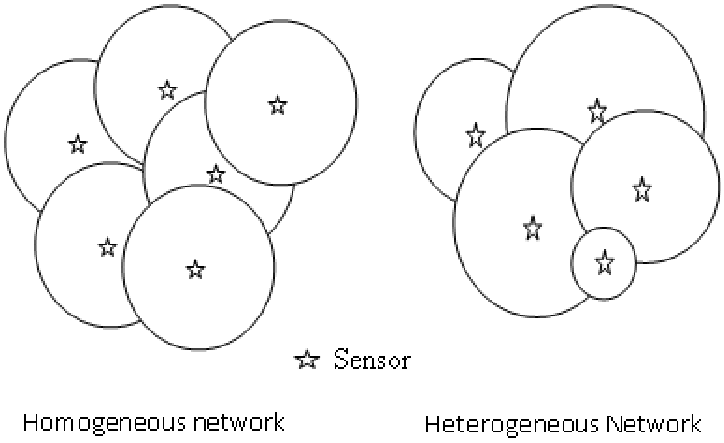

Figure 2.3.: Homogeneous and Heterogeneous Network

Dimensional coverage

WSN is employed in either two dimensional or three dimensional space. It is much easier to formulate protocols for two dimension as compared to three dimensional space. The paper [82] was one of the first to describe 3-D space.The authors assume that the sensor's coverage ranges are in the form of a sphere.

Sensing coverage and network connectivity are two of the most fundamental problems in WSNs. Finding an optimal node deployment strategy that would reduce computation and communication overhead, minimize cost, be resilient to node failures, and provide a high degree of coverage with network connectivity is extremely challenging.

2.5.2. Approaches to Coverage in WSN

Art Gallery problem

The Art Gallery problem in computational geometry says that a point x is visible by another point y (guard) if the entire straight line from x to y is within the polygon (area) [83]. Applied to coverage in WSN, a guard represents a sensor and the polygon depicts the area to be sensed. If the entire polygon is seen by any guard then it is said to be covered.Howard et al[84] state that the art gallery problem can be used as a base for a coverage algorithm only when the shape of the area to be sensed is known before deployment. It would usually only be used with the deterministic placement of the sensor nodes in WSN.

Voronoi diagram and Delaunay triangulation

The Voronoi diagram for a WSN is a diagram of boundaries around each sensor such that every point within a sensor's boundary is closer to that sensor than any other sensor in the

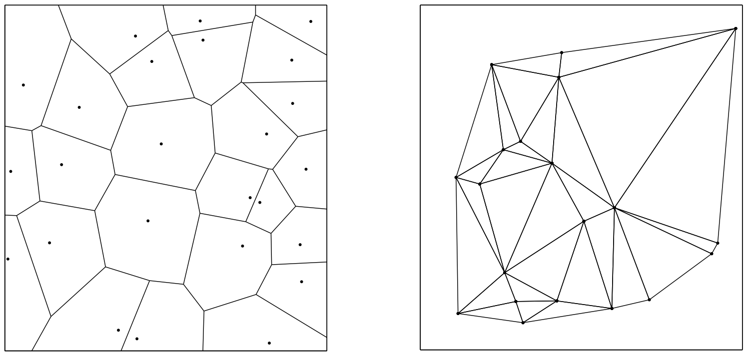

Figure 2.4.: Voronoi Diagram and Delaunay triangulation

network. A formal definition of the Voronoi diagram [83]is

Let $P = \{p_1, p_2, ..., p_n\}$ be a set of points in a plane

A Voronoi region $V(p_i)$ is the set of points that are as close to p_i as any other point as mentioned in the equation 2.1

$$V(p_i) = \{x : |p_i - x| \leq |p_j - x| \, for all \, j \neq \text{i}\} \tag{2.1}$$

The authors[85] have solved two coverage problems using Voronoi diagrams. The Delaunay triangulation is the triangulation of an area such that there are no points in any triangle which are located within the circumscribed circle of any other triangle in the area. It can be built from a Voronoi diagram by drawing edges that connect the sensors which border one another. The Delaunay triangulation is used to find the nodes which are at the shortest distance. Neither the Delaunay triangulation nor the Voronoi diagram can be constructed with localized algorithms [86]. Figure 2.4 shows the two diagrams.

The authors[81] tackles the coverage problem by using Voronoi diagrams generated with delaunay triangulation.

Disjoint sets

When the nodes are densely randomly deployed, it may happen that many nodes are placed in one single region which are not actually needed for the required coverage of that particular region. In such cases, the unneeded sensors can be put to sleep, thus conserving the overall energy of the network. One way to accomplish this goal is to divide the sensors into groups or sets.Each set should be able to cover the area to be sensed.The disjoint set cover is a subset of the sensors that is capable enough of covering the entire area by itself [87]. Each set cover is activated and put to sleep in turn so as to conserve the overall energy of the network. Zhao et al[88] discuss the shut off and turn on of sensors in the network.

2.6. Connectivity in WSN

In addition to coverage it is important for a WSN to maintain the connectivity. Connectivity is defined as the ability of the sensor nodes to reach the sink. If there is no route available from a sensor node to the sink then the data collected by that node can not be processed by the sink. Each sensing node has two types of ranges i.e. sensing range and transmission range. The area which the sensor is capable to sense is called as the sensing range. The transmission range is the area up to which the node can transmit the data to other node or to the sink. The subsequent sub sections discuss the network connectivity issues and some proposed solutions to connectivity problems in WSN.

2.6.1. Connectivity Issues in WSN

The node sends the data to another node or the sink within its transmission range. The connectivity of a network is maximum when there is a path between all the nodes of the WSN to the sink. This path can be direct or through multiple hops. If the link between the two nodes or between the node and the sink breaks, then the connectivity of the network is broken. There are many reasons a link between nodes can fail, like sparse amount of nodes, physical damage to nodes, Energy depletion of nodes, Security threat/denial of service, Environmental changes and Mobility of nodes.

Sparse amount of nodes

When the nodes are randomly deployed, there may be very few nodes in a region of the area to be sensed. In such regions, due to limited transmission energy of the nodes, the connectivity of the network is effected and leaves the network with communication holes[89].

Physical damage to nodes

There can be a physical damage to the sensor nodes either due to environmental conditions or through can be crushed by an animal, enemy soldiers etc. Depending upon the region where the nodes are deployed, the nodes can be damaged by different ways. Once any node gets damaged, its link is broken with the rest of the network. It may also be a part of a multi hop route of one node to the sink, thus affecting the connectivity of the entire network.

Network security threat

The nodes are always susceptible to security threats, specially when the WSN is used in highly confidential regions like military areas etc. In such cases, when any nodes gets hacked, the connectivity of the entire network is affected [90, 91].

Energy depletion of nodes

The sensor nodes have limited source of energy. Generally the WSN are employed in regions where the nodes are unattended for very long duration. Thus, there are no options of recharging the battery of the sensor nodes. A lot of energy is consumed in transmitting the data to other nodes or the sink. The energy of the nodes get depleted fast. Various algorithms have been proposed by the researchers for energy efficient transmission [92, 93, 94]. But when the nodes get die out due to energy depletion, the connectivity of the network is seriously affected.

Environmental changes

The environment in which WSNs operate may change due to certain circumstances. Radio connection is very sensitive to rain and growing plants, which consequently changes the network topology [95]. When the WSN is employed in some cultivation fields or in forest, the plants grow up after some time and prevent the nodes to connect to each other leading to in-connectivity in the network.

Mobile nodes

Nodes may be mobile as required by the application or may be displaced by human and animals. It may also be floating with running water if deployed in the sea. Mobility of nodes causes a great impact on the network connectivity [96]. When a node selects a routing path and the nodes in the routing path change their locations, it will change the connectivity between source and destination. In an another case, if a data packet is long and due to mobility the node changes its current location, part of the data may be lost at the receiving node.

Coverage and connectivity together can be treated as a measure of quality of service in a sensor network; it tells us how well each point in the region is covered and how accurate is the information gathered by the nodes. Therefore, maximizing coverage as well as maintaining network connectivity using the resource constrained nodes is a non-trivial problem.

2.6.2. Solution to Connectivity in WSN

A solution to maintain or to re-establish network connectivity is to remove the nodes which are creating connectivity problems and deploy more nodes in the sensor field. But generally the WSN are employed in regions where it's not possible to replace a node or deploy more nodes to maintain the connectivity. Thus efficient routing protocols and WSN algorithms should be used to curb down the issues of connectivity.

Graph models

A WSN is often represented by a graph in which vertices represent the sensors and a directed corresponds to the communication link between one node to another. The communication link signifies that one node can transfer the data to the other. Since it's a directed graph, a directed edge between node A and node B means that node A can send the data to node B but the vice versa may not be true. But if all nodes have equal transmission ranges as in a homogeneous network, then the graph becomes undirected. A network is called connected if this associated graph is connected[97, 98] . It is sometimes useful to consider stronger forms of connectivity, like k-connectivity, in which the network remains connected even if $k-1$ nodes are removed. A k-connected network $(k > 2)$ has better fault-tolerance than 1-connected[99]. Random graphs may also be applied to model communication networks to highlight their randomness. Mathematically, a random graph is generated by a stochastic process [100].

2.7. Multiple Sinks

The issues of coverage and connectivity in WSN have already been discussed. Some of the above mentioned issues are resolved by the use of multiple sinks. In a WSN, the fundamental question is to have the data routed over single hop or multiple hops. This question is answered by considering the answer to the question that the data needs to be sent over a longer or a shorter hop. Short-hop routing leads to reduced energy consumption and higher signal-to-interference ratios[101]. The less but longer hops lead to more energy consumption but less Signal to Interference Ratio (SIR). Research paper [102] it's proved that single-hop transmission is more efficient, when power consumption of real wireless sensor node's transceivers are taken into account. So, this leads to four types of networks:-

1. Single hop single sink routing (SH-SS)
2. Single hop multiple sink routing (SH - MS)
3. Multiple hops single sink routing (MH – SS)
4. Multiple hops multiple sink routing (MH – MS)

The first scenario is the most elementary one with direct transmission. The LEACH protocol which is better than the direct transmission, deals with single hop single sink clustering protocol [3].

A lot of research is done for the third scenario [47, 103, 104]. The modification of LEACH protocol having multiple hops instead of single hops[47]. In [103] compares the single and multiple hop routing.

Table 2.6.: Related work in multiple sinks in WSN

Research Paper	Hops	Sinks	Mobility of the Sink	Sink Positioning	Network Restructuring
[110]	Multiple	Multiple	Yes	No	No
[109]	Multiple	Multiple	Yes	Yes	No
[19]	Single	Multiple	No	Yes	No
[111]	Single	Multiple	Yes	Yes	No
[4]	Multiple	Single	No	Yes	No
[3]	Single	Single	No	No	No
[47, 103, 104, 106]	Multiple	Single	No	No	No
[16]	Multiple	Multiple	No	No	No

In [16], a multi hop protocol spends most of its energy for relaying data packets so the concept of multi hop multi sink WSN is discussed. The concept of particle swarm optimization is used.

The current research work deals with the second scenario i.e. Single hop multiple sink. The biggest disadvantage of single sink is that certain sensors near the sink or on critical paths consume energy much faster than other nodes [105]. Thus the current work uses the advantage of having multiple sinks. Multiple sinks ensure shorter hops and thus the 18 advantages as discussed in the paper [101] are also achieved. Multiple hops are generally used to reduce the hop distance [106]. But if multiple sinks are used, the hop distance automatically reduces. Thus, the research deals with single hop and thus avoiding the drawbacks of having multiple hops.

Network lifetime (NL) is a critical metric in the design of energy-constrained WSN[107]. The basic aim of the researchers is to minimize the energy consumption and at the same time increase the network lifetime. The authors in the paper [108, 109] deal with mobile multiple sinks. Data dissemination to multiple mobile sinks consumes a lot of energy [110]. Many papers [108, 16] have concentrated on positioning of the sink to have optimal energy consumption. The current research work talk about the random deployment of the sink thus saving power in determining the position of the sink. Then the network is re-structured to have balanced energy consumption amongst all the sinks. Table 2.6 summarizes the related work done in the field of single/ multiple hops, single/multiple sinks and moving/stationary nodes.

Depeding upon the requirements of the application where the WSN is going to be used, the factors mentioned in the Table 2.6 are chosen. If the network is very large and cost effectice then single sink multiple hop network is used. If the cost can be increased to make the network lifetime better, then multiple sinks can be used. The mobile sinks increase the network lifetime to a great extent [110]but finding out the location of the sink makes the network computationally complex[112]. The authors in the research paper [111]talk about lexicographically optimal commodity lifetime (LOCL) routing problem to position the multiple sinks at optimal positions in order to increase the network lifetime. Network restructuring is changing the neighbor nodes connected to a sink depending upon the energy consumption by that sink. In the current research work the number of sensors connected to any sink is changed if the energy consumption by the sink is more than the threshold. To have balanced energy consumption amongst all the nodes, the entire network is re-structured.

2.7.1. Multiple Sinks Placement

Oymen et al [113]proposed an algorithm for finding the position of sinks using k-means clustering. The cluster centroids are chosen as the optimal position of sinks. The sink placement problem is NP-complete [114], and finding the best position of sink is very hard. The authors in the research paper [115] introduce some sink placement strategies and also discuss their pros and cons. The Geographic Sink Placement (GSP)[116] strategy places the sinks at center of sector of a circle. In Intelligent Sink Placement (ISP), the possible sink locations are determined by sampling all possible regions. All combinations of these candidate locations are found out, depending on the number of sinks. Then the optimal location of the sink is calculated. ISP is found to be optimal buts it is computationally expensive. Another algorithm, called Genetic Algorithm-based sink placement (GASP) [117]provides a good heuristic based on Genetic Algorithm for optimal sink placement. In [118] the authors have proposed an algorithm for sink placement using linear programming. The problem of finding the optimal number and the position of the sink nodes resembles some classical problems like problem of plant location [119], problem of warehouse location [120], and the concentrator location problem (CLP) [121]. But, there are many differences between these problems and the multiple sink location problems as stated in paper [113]. First of all the sinks can be placed anywhere in the network and secondly the transmission are done in multi hops as compared to direct in the traditional problems.

2.8. Radio Model

In a wireless channel, the electromagnetic wave propagation can be modeled as falling off as a power law function of the distance between the transmitter and receiver. In addition, if there is no direct, line-of-sight path between the transmitter and the receiver, the electromagnetic wave will bounce off objects in the environment and arrive at the receiver from different paths at different times. This causes multipath fading, which again can be roughly modeled as a power law function of the distance between the transmitter and receiver. No matter which model is used (direct line-of-sight or multipath fading), the received power decreases as the distance between the transmitter and receiver increases [18].

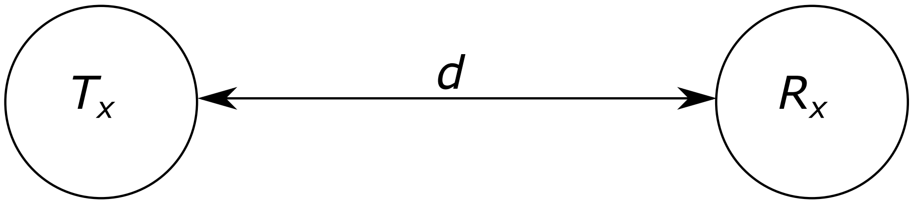

Figure 2.5.: Radio Model

For the simulation performed in this thesis, both the free space model and the multipath fading model were used, depending on the distance between the transmitter and receiver, as defined by the channel propagation model in [18]. If the distance between the transmitter and receiver is less than a certain transmission range (d_0), the Friss free space model is used $(d^2 attenuation)$, and if the distance is greater than transmission range (d_0), the two-ray ground propagation model is used $(d^4 attenuation)$. The transmission range is depends upon the different wireless technologies such as Bluetooth [122], Zigbee [123], and GSM [124] etc. In most of the cases sensor network uses Zigbee technologies because low power requirement, easy protocol stack and long distance transmission. The transmission range point is defined as follows in Eqn: 2.2

$$d_0 = \frac{4\pi\sqrt{L}h_t h_r}{\lambda} \tag{2.2}$$

where

$L \geq 1$ is the system loss factor not related to propagation,

h_r is the height of the receiving antenna above ground,

h_t is the height of the transmitting antenna above ground,

λ is the wavelength of the carrier signal, and

d_0 or R_t is the transmission range.

A great deal of work is going on energy consumption of radio models in Figure 2.5 and eq. 2.3 to 2.8. Different assumptions made in radio models have different advantages. In our work, we consider a simple model where radio dissipates $E_{elec} = 50nJ/bit$ to run transceiver circuitry and $\varepsilon_{fs} = 100pJ/bit/m^2$ for transmitter amplifier, so as to achieve acceptable SNR ratio[3].

To transmit information as given in eq. 2.5 transmitter expends:

$$E_{Tx}(k, d_{ij}) = E_{Tx-elec}(k) + E_{Tx-fs}(k, d_{ij}) \tag{2.3}$$

$$E_{Tx}(k, d_{ij}) = E_{elec} * k + \varepsilon_{amp} * k * (d_{ij})^n \tag{2.4}$$

$$E_{Tx}(k,d_{ij}) = E_{elec} * k + \varepsilon_{fs} * k * (d_{ij})^2; d_{ij} < d_0 \tag{2.5}$$

$$E_{Tx}(k,d_{ij}) = E_{elec} * k + \varepsilon_{mp} * k * (d_{ij})^4; d_{ij} > d_0 \tag{2.6}$$

$$E_{Rx}(k) = E_{Rx-elec}(k) \tag{2.7}$$

$$E_{Rx}(k) = E_{elec} * k \tag{2.8}$$

Where
E_{elec} Energy dissipation in electronic circuitry in Transmitter and Receiver,
k number of bits,
d_{ij} is the distance between i^{th} node to j^{th} node,
n path loss exponent,
ε_{amp} is proportionality constant,
ε_{fs} is proportionality constant in free space,
ε_{mp} is proportionality constant in multi path, and
d_0 transmission range.

Using these parameters receiving message is not a low cost operation. So protocols used in network should try to minimize not only the distance between two nodes but should also minimize number of transmissions and receptions for each message. Path loss exponent (n) depends upon different environments [125] and are shown in Table 2.7

Table 2.7.: Path loss exponents (n) for different environments

Environment	Path loss exponent
Free Space	2
Urban area cellular radio	2.7 to 3.5
Shadowed urban cellular radio	3 to 5
In building line-of-sight	1.6 to 1.8
Obstructed in building	4 to 6
Obstructed in factories	2 to 3

2.9. Summary

Having discussed the different types of routing protocols, design parameters of WSN, Coverage issues and the approaches to coverage , connectivity issues in WSN and their solutions and finally multiple sink, the author proposes his work in different chapters of the thesis.

The main objective is to minimize the energy consumption, maximize the coverage and connectivity and the optimization of the network lifetime. The next chapter deals with a novel algorithm which employs clustering in a single sink WSN so as to have the minimal energy consumption. If the energy consumption is minimized the network lifetime is automatically increased. A WSN is designed to be deployed to sense multiple physical attributes in an environment and produce an alarm if any unexpected condition is predicted.

Part I.

Single Sink Wireless Sensor Network

3. Energy Efficiency in Single Sink

3.1. Introduction

This chapter focuses on the design of an optimum energy multi-hop network through clustering. As discussed in the previous chapter, energy is a major constraint in designing a WSN. WSNs are low power battery driven networks, hence efficient energy utilization is needed. This chapter focuses on designing an optimum energy multi-hop network. For this purpose the concept of merging is proposed in which if two sensing nodes are falling in the same sensing range then they will work in sleep and active mode as a result of which less energy will be used. Nodes are arranged in the form of clusters and each cluster has a cluster head (CH). All the nodes communicate with base station using multiple hops via CHs. The energy is computed in terms of number of rounds of transmission and reception of information. Along with the efficient energy usage, the proposed merging algorithm helps in finding out the lifetime of network (in number of rounds).

The section 3.2 talks about clustering. The section 3.3 deals with a system Wireless Environmental Monitoring System (WEMS) which is a WSN used to monitor the environment and it also warns for a mishappening occurring in future. The section 3.4 talks about the clustering algorithm using sensing range as the parameter for CH selection.

3.2. Clustering

Wireless Sensor Networks have unlimited applications all around us and these networks will enable the reliable monitoring for various environments for both civil as well as military purposes. WSNs are low power battery driven networks, hence efficient energy utilization is needed. Clustering is a useful mechanism in WSNs which helps to cope with scalability and data transmission problems. Different clustering are depicted in the Figure 3.1.

A flat network mostly follows data centric protocols. In such a network all the nodes are assigned an equal role and functionality. The desired data is sent out to the network through multi-hop routes. To eliminate many redundant transmissions through the network, flat protocols focus on how to route based on the application queries. Nodes only transmit the valuable data which match the query attributes. In many cases, flat protocols result

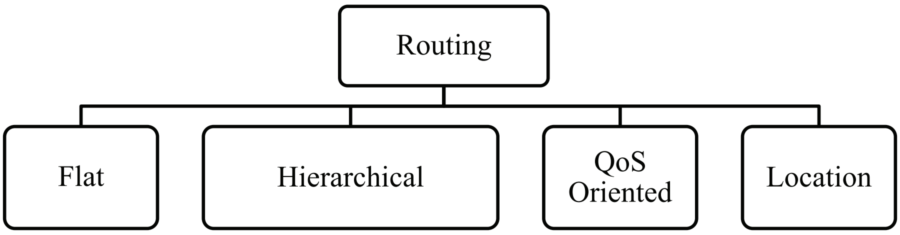

Figure 3.1.: Clustering

in more complicated routing because of the large scale and dynamic network topology of WSNs. Sensor protocols for information via negotiation (SPIN) [38] and directed diffusion (DD) [40] protocols are important flat protocols which motivated the design of many other protocols that follow similar concepts.

Clustering and prediction techniques, which exploit spatial and temporal correlation among the sensor data, provide opportunities for reducing the energy consumption of continuous sensor data collection [126]. The research paper [127] deals with overlapping clustering. Overlap clusters are those in which a node may belong to more than one cluster, in contrast with the traditional clustering algorithms, in which each node belongs to only one cluster. The paper shows that the overlapping multi - hop clustering problem is NP-hard problem so the k-hop Overlapping Clustering Algorithm (KOCA) is proposed which is a randomized distributed algorithm for solving it. The authors [128],deal with single hop clustering in which a lot of energy is required by the nodes. The paper deals with a very useful application of WSN i.e. wireless body area sensor networks (WBASN) used in healthcare applications.Sensor nodes are attached to the human body to monitor vital signs such as body temperature, activity or heart-rate. To avoid collisions with nearby transmitters, a clear channel assessment algorithm based on standard listen-before-transmit (LBT) is used.

ACE [129] successfully distributes clusters uniformly over the network but suffers from its unawareness of residual energy in cluster-heads candidates, which results in electing a cluster-head with low energy level. The other disadvantage of ACE strictly draws a line between nodes that can be a cluster-heads and the ones who can't. In some cases this assumption may be unrealistic, especially when all the nodes within a cluster have low power resources.

Hierarchical clustering is a clustering scheme in WSN[130] where nodes play different roles, such as CHs and cluster members. The higher level nodes, CHs, manage the lower level nodes (cluster members) which are grouped under them. The cluster based algorithms are used to partition the sensor nodes into subgroups for task subdivision or energy management. Each CH collects the data from its cluster members within its cluster, aggregates the data, and then transmits the aggregated data to the base station. All the communication to (from) each sensor node is carried out through its corresponding CH. Since the CH is closer to the node compared to the BS, the energy is automatically saved. CHs facilitate various aggregation techniques to fuse data from sensors to minimize the amount of data to be sent to the sink. The aim of all of the hierarchical routing protocols is to select the best CH and clustering the nodes so as to save energy. Since the CHs collect, aggregate, and transmit data over longer distances to the sink, they consume more energy as compared to the other cluster members. The hierarchical clustering protocol need to re-cluster and reselect CHs periodically in order to distribute the load uniformly in the entire sensor network.

Cluster formation is one of the most important problems in sensor network applications and can drastically affect the network's energy dissipation during communication. Clustering is performed by assigning each sensor node to a specific CH. When a sensor with sufficient battery and computational power detects (with a high Signal-to-Noise Ratio: SNR) signals of interest, it volunteers to act as a CH. This is a simple method, because no explicit leader (CH) election is required and, hence, no excessive message exchanges are incurred. However, selecting the CH in this way is not easy in different environments which may have different characteristics such as error rate, SNR, throughput and so on [131, 132].

Clustered sensor networks can be classified into two broad types; homogeneous and heterogeneous sensor networks [133]. In homogeneous networks all the sensor nodes are similar in terms of battery energy and hardware complexity, whereas, in a heterogeneous sensor network, two or more different types of nodes with different battery energy and functionality are used [134].

Compared with flat protocols, hierarchical protocols offer a more feasible solution to handle large-scale networks with their enhancements to better share limited wireless channel bandwidth, balancing node energy consumption and reduce communication expense [135, 136].

The advantages and disadvantages of the two kinds of hierarchical routing protocols are summarized are as follows:-

1. Random-selected-CH protocols. Although randomly selected-CH protocols are more flexibile and tolerant, these approaches have three main disadvantages. Firstly, the randomly selected CH may have a higher communication cost because it has no knowledge of intra-cluster or inter-cluster communication. Secondly, if periodic CH rotation is used to reduce the effect of CH random selection, the re-selection itself uses extra energy to re-build clusters. Periodic CH rotation also leads to an uneven wave of performance due to the nonstop change. Thirdly, the random selection cannot guarantee good protocol performance. In other words, the best arrangement and the worst arrangement have an equal chance to be used in the network (LEACH) [3]

2. Well-selected-CH protocols. The well-selected-CH protocols can provide better clus-

ter quality, but they usually have a more complex scheme and higher overhead to optimize the CH selection and cluster formation. Some approaches use the sink to help choose CHs by frequently collecting information from nodes. However, the sink performing the algorithm introduces another issue that increases communication cost between the nodes and the sink because they need to frequently exchange administrative information. Other researchers have to try to use the optimization algorithms to distinguish the roles of nodes. But there may not be enough fault tolerance in these schemes because any change to the network may cause the entire network to update information and perform re-clustering (LEACH C)s [51]

To perform clustering, the data containing the exact position of the nodes is usually used for calculating the distance between sensor nodes. But the location of the sensing nodes may not always be available due to Global Positioning System (GPS) failures or may not be practical due to the involvement of a large cost. Hence, alternatively, Received Signal Strength (RSS) or RSS Indicator (RSSI) is used to calculate the distance between the nodes [137].

Given a sensor network $N = S_1, S_2, \ldots, S_n$of sensing nodes and an integer value k, the clustering problem is to define a mapping

$$f : N \rightarrow \{1,2,...,k\}$$

where each S_i is assigned to one cluster $K_j, 1 \leq j \leq k.$ is the total number of clusters in the network. A Cluster, K_j, contains precisely those sensing nodes mapped to it; that is

$$K_j = \{S_i | f(S_i) = K_j,\ 1 \leq i \leq n\, and\, S_i \in N\}$$

When clustering is applied in a real world scenario like a WSN, following problems or limitations may occur:-

1. Outlier handling: it may happen that a node doesn't belong to any cluster and it resides in a cluster of its own. This makes the clustering inefficient

2. Dynamic data: the sensed data changes continuously and the algorithm should be able to handle such dynamic data

3. Interpretation: interpreting the semantic meaning of each cluster may be difficult. With classification, the labelling of the classes is known ahead of time. However, in clustering, the exact meaning of each cluster may not be obvious at real run time. Here is where a domain expert is needed to assign a label or interpretation for each cluster.

4. No one correct answer: the exact number of clusters required is not easy to determine. A domain may be required.

3.3. Wireless Environment Monitoring System

The Wireless Environment Monitoring System (WEMS) monitors critical environmental conditions, such as temperature, humidity, intrusion, and smoke[4]. When the sensed attribute goes out of the range of a threshold, the system will notify using an alarm. The sensor network measures the attribute which it asked to, like temperature, pressure etc, at a particular point and pass on this information to the application processor. The user can get his desired information directly from the application processor in the form of a topographic map. The end user might be interested in extracting a variety of topographic information about the metadata in the network. He might want to know the boundaries of all regions where the temperature exceeds a certain threshold, or just the numbers of disjoint regions where temperature exceeds a threshold, or might want the complete topographic map of the terrain where contours correspond to temperature levels.

The proposed network is capable of monitoring more than one physical attribute as opposed to single attribute detection. This chapter considers about mobile sensor nodes, whose position is calculated by the CH through triangulation method. The benefit of data aggregation can be maximized by implementing it at each CH till the Base Station (BS) or the sink. The data need not be concealed or made private, so the overhead of encryption and decryption is avoided. The BS uses distributed computing to make the processing of the sensed data fast and more efficient. The protocol discussed is bi-directional and a hybrid protocol, which is a combination of reactive and proactive approaches.

Data aggregation (DA) collects the most critical data from the sensors and make it available to the BS in an energy efficient manner with minimum data latency [138]. The advantage of it is that the data transmissions are minimized as the data is first aggregated at a node and then it is send to the BS.

3.3.1. Architecture of the Proposed System

The protocol is bi-directional and hybrid. In addition to getting the value at regular intervals, it can also react to any unexpected or emergency condition. The application processor can ask the sensing unit to send the data at a particular instant and it can even change the time interval at which data is being transmitted by the sensing unit. Thus the time critical data reaches the user on time without any delay.

The protocol facilitates the sensing of more than one parameter by the sensor network. The sensor node passes the complete data to the BS. It sends the value of the attribute along with the sensed attribute, so that the BS knows which parameter is sensed by the sensor node.

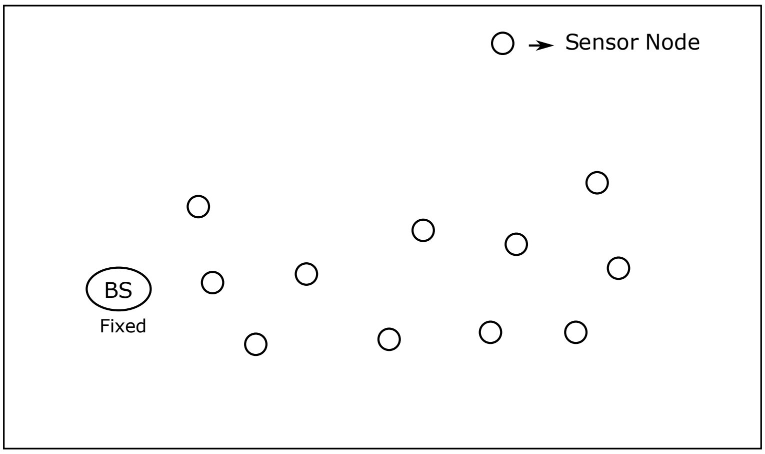

Figure 3.2.: step 1 of cluster head selection

3.3.2. Cluster Head Selection

The first step in transmitting the data in the WEMS is the selection of CHs and the creation of the clusters. The following steps are followed:

Step 1: Sensors are randomly deployed on the sensor field and the BS is fixed to transmit the sensed information to the application processor as shown in the Figure 3.2.

Step 2: BS transmits the Cluster Head Selection message (CHS_msg) to the all the sensors within its transmission range. It assigns a hop_ID of 1 to all the sensor nodes. In the Figure 3.3 the two nodes coming within the transmission range of BS are A and B. The CHS_msg consists of the following parameters:

1. *hop_ID*, the number of hops from the current node to the BS.

2. *msg_type*, the type of message send by the BS to the sensor network. It can be a CH selection message or it can be a query processing message.

3. *ch_ID*, the id of the previous CH or the BS.

Step 3: In Figure 3.4, Sensor node A only send the *CHS_msg* to C and D, so in this case A become CH. After each CH transmit the message to child sensors in the network and then decide the CH selection.

Step 4: The transmission range of A sends the *CHS_msg* to n number of nodes than the selection criteria of CH as follows

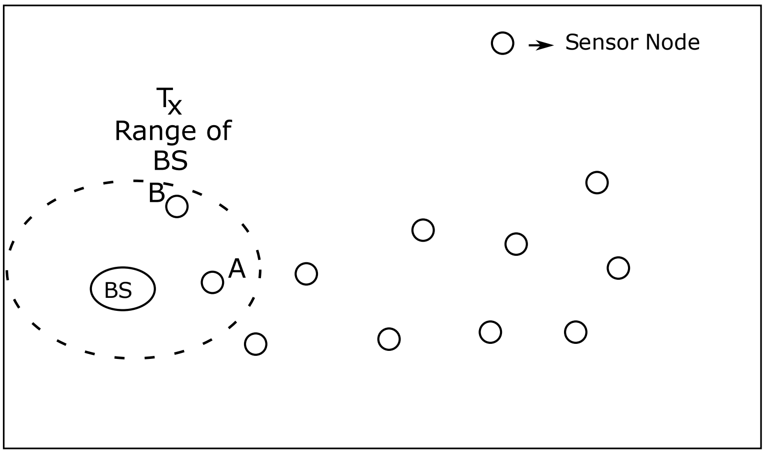

Figure 3.3.: step 2 of cluster head selection

Algorithm 3.1 Cluster Head Selection

$if\,(!CH)$
$\{$
  $if\,(!CH\,in\,last\,n-1\,rounds)$
  $\{$
    $if\,(!Receive\,CHS\_msg\,in\,current\,round)$
    $\{$
      $Calculate\,R_n = generate\,random\,between\,0\,to\,1;$
      $if\,(R_n < T(n))$
      $\{$
        $Nodes\,become\,CH;$
        $BroadcastCHS\_msg\,to\,child\,sensor;$
      $\}$
    $\}$
  $\}$
$\}$

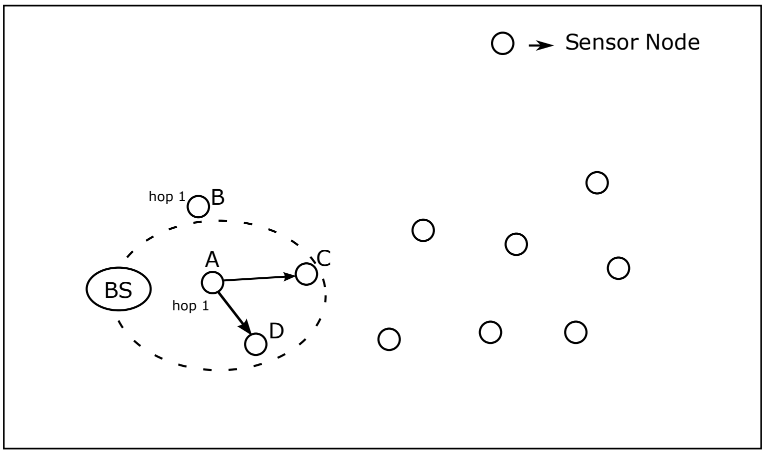

Figure 3.4.: step 3 of cluster head selection

Initially, when clusters are formed, each node decides whether it's the CH for the current round or not. This decision is based on the percentage of the CH for the network (determined a priori). This decision is based on random generator value between 0 to 1, if the number is less than a threshold $T(n)$, node become CH for the current round. The $T(n)$ = Threshold value can be calculated by given formula [139].

$$T(n) = \frac{p}{1 - p * (r mod(1/p))} \quad n \in G \tag{3.1}$$

where p = the desired percentage of cluster heads, r = the current round, and G is the set of nodes that have not been CH in the last 1/p rounds.

Step 5: Once all the CHs are formed in the given network and on the basis of *hop_ID*, decide the routing path having minimum number of hopsin as shown in the Figure 3.5. In the given network node A, C, F and as CH and all other nodes are needed $hop_id + 1$hop to send the data from source node to sink node.

In this way the CHs are selected and cluster formation is complete. The black circles in the diagram represent the CHs. The sensors know their CHs, the ones which have minimum hop_id. E.g. if a sensor node is in the transmission range of two CHs, it received the *CHS_msg* from two CHs as $(3, CHS_msg, CH_4)$and $(5, CHS\_msg, CH_6)$, where 3 is the *hop_id*,*CHS_msg* is the message type and CH_4 is the CH id. Then the sensor knows that its CH is $CH_4$and not CH_6 because the number of hops required sending the data from the node to the sink in the former case is 3 and in the latter casing its 5.

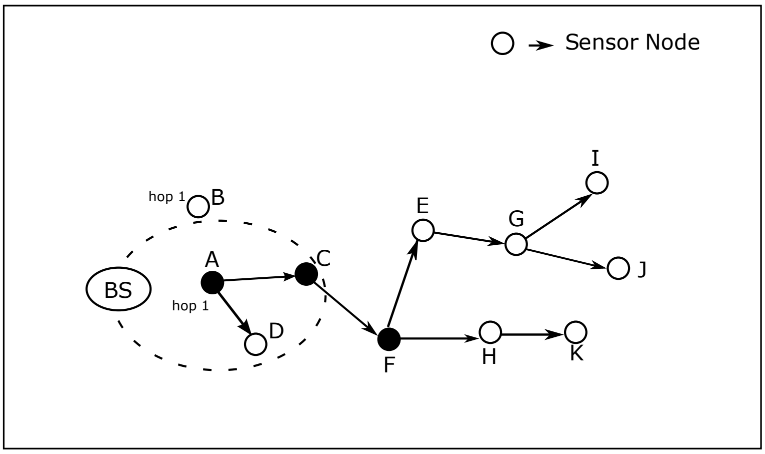

Figure 3.5.: Cluster head selected

3.3.3. Data Transmission from the Sensor to the Sink

The reactive part of the protocol deals with the transfer of data from the sensor node to the BS whenever there is some unexpected behavior sensed by the sensor. The proactive part of the protocol deals with transmission of data from the sensor to the BS at regular time intervals. This time interval is changed by the application processor as and when required. The sensors send their unique identification number (S_{id}), the value of the parameter to be measured (V_iA), and the time of measurement (t) to the CH selected earlier. The data transmitted would be of the form (S_id, V_iA, t)

The sensor sends the data to the CH assigned to it. The CH sends the information to the BS through the minimum number of hops as designated by the *hop_id* of the CH. The CH and consequently the BS are aware of the position of the mobile transceivers through the triangulation method. Using the signal strength the CH performs triangulation to determine the exact location of the sensor. And correspondingly it adds the coordinates of the sensor in the message to be sent to the next CH (next hop) or to the BS. The format of the information is sent.

$$(S_id, V_iA, t, (X_i, Y_i), CH_{id}) \tag{3.2}$$

Where CH_{id} is the CH id. The BS would use the above information to process any type of application. In this task it may generate spatial and temporal data, process the desired parameters and predict events like cyclone, Tsunami etc and the affected areas.

3.3.4. Data Aggregation

As discussed in section 3.3, the proposed protocol deals with the sensing of multiple attributes by the sensor network. Jiao Zhang [140] talks about attribute aware data aggregation, which deals with multiple attributes being sensed by the network and their values being aggregated.

In the case of WEMS, a physical quantity is sensed. In real world, the environmental factors (e.g., temperature, humidity) change continuously instead of flipping over along the space field. According to this natural phenomenon, there exists kind of correlation in the data gathered from natural environment. Hence in WEMS, data aggregation can be done by exploiting the correlation among the data. This would save time and energy. One way to exploit correlation is a linear transform in which the statistically dependent data will be mapped into a set of more independent coefficients and then compressed and transmitted. 3D-DCT (Discrete Cosine Transform) algorithm is used to exploit the spatial-temporal correlation in a regularly deployed network to achieve significant aggregation performance [141]. The DCT transform function for N data sequence, f, is defined in [142].

Considering an example, suppose there are 10 sensor nodes connected to one CH and each node is sensing the temperature. In the case of WSN, each node send the sensed data to the base station, either in single or multiple hops. If there are 2 clusters, then there would be total 20 nodes sending their data to their respective CHs and the CHs in turn sending the data to the sink. So, the base station has a huge amount of data which is of the order of 20 *packet_size. The nodes close to each other, send the redundant data to the CH creating an unnecessary traffic. To avoid this, the data is aggregated at the CHs. So, 10*packet_size data reaching the CH is reduced by sending only the minimum, maximum and the average value of the temperatures. So, instead of 20*packet size data reaching the sink, there is only ((3*2)*packet_size) data reaching the sink. Thus, the data aggregation reduces the communicated data to a considerable amount.

3.3.5. Simulation and Results

The network is simulated in OMNet++. OMNeT++ is a component-based, modular and open-architecture discrete event simulation framework. The simulation results show that the proposed protocol is good in energy conservation by effectively selecting the CHs. The data can flow in both the directions with the minimal loss and time delay. It has been observed that there can be an effective hybrid bidirecional protocol for the transmission of sensed data in a WSN. Sensing multiple attributes, increases the complexity of the system and the size of the message passed by the BS to the sensor network also becomes more as compared to a single sensed attribute. The WEMS can analyze the parameters of the environment and can produce alarm in adverse conditions.

From the simulation results of a given network in Table 3.1, it is evident that with the help of data aggregation technique the Network Lifetime (Rounds) and Packet Delivery Rate is increased. Due to data aggregation mechanism the size of packet reduces so that the

Table 3.1.: Simulation Results

Wireless Environmental Monitoring System (WEMS) Using Data Aggregation			
	Rounds		Packet Delivery
	First Node Die	Network Lifetime	Efficiency (%)
Without Data Aggregation	50	60	47
Data Aggregation near BS	63	94	64
Data Aggregation all Hops	120	147	92

successful packet delivery at BS increases. These are the primitive level results and some strong conclusion can be drawn from the subsequent section of this chapter. From Table 3.1, it can be observed that when data aggregation is used near the BS or at all the Hops, the number of rounds for the First Node to Die in the network increases by 13 and 70 rounds as compared to the case that no data aggregation mechanism is used. In the same way, Network Lifetime and Packet Delivery Efficiency is also observed to improve when such data aggregation mechanisms are used in the considered network.

3.4. Clustering using Sensing Range

To modify the clustering process further to make it even more energy efficient network we propose the concept of merging in a multi-hop network in which if two sensing nodes are falling in the same sensing range then they will work in sleep and active mode alternatively as a result of which less energy will be required [9] (Detailed explanation can be found later in Section 3.3.3). The energy used in network is computed in number of rounds of transmission and reception of information. Along with the efficient energy usage, our proposed merging algorithm helps in finding out the lifetime of network (in number of rounds).The notation used in Merging algorithm are depicted in Table 3.2

The network used for the current research work has the following properties:

1. The nodes are homogeneous having equal initial energy (E_0) of 1 joule.
2. The nodes transmit the data to the BS in multiple hops.
3. The hops are determined based on the distance of the node from the BS.
4. Considering $R_t = 150m$ as the transmission range and $R_s = 25m$ as the sensing range, we have considered $R_t \geq 2R_s$to be a valid assumption.
5. The N nodes are randomly deployed.
6. After random deployment sensor nodes are stationary.

Table 3.2.: Notations used in Merging

Notations	Meaning
N	Total nodes in the network
E_0	Initial node energy $1J$
n_0	Node ID of BS
n_i	Node ID ofi^{th}node
K	Packet size (no. of bits)
E_{th}	Threshold energy value at which the CH dies
E_{DA}	Data aggregation energy
R_s	Sensing Range
R_t	Transmitting Range
N_i	Set of neighboring nodes of i^{th}node
CH_{id}	Cluster head ID's for each nodes
E_r	Residual Energy < E_0
D_{ij}	Distance between i^{th}node and j^{th}node
H_{id}	Number of hops to reach the BS

7. BS (n_0) is fixed and deployed somewhere in the middle of the network (250,250).

8. The nodes are proactive.

They transfer data to the BS in periodic intervals. To build a network we start from the BS and connect all nodes lying in its transmitting range. Then these connected nodes connect to the sensor nodes lying in their transmitting range. This process continues till all the nodes are connected. After all the connections are made the routing table of the network is calculated. The routing table tells us about the path of every node and number of hops. To calculate the life time or the energy of the network we have used first order radio model. By calculating the energy required in transmitting and receiving we can estimate number of rounds a network can last.

3.4.1. Cluster Head Determination

Step 1: The nodes in the transmitting range of the BS are connected to the BS.

Step 2: The sensors deployed within the transmitting range are connected to the already connected node. Any node connected to more than one node is called a CH.

Step 3: BS transmits the CHS_msg to all the sensors within its transmission range. It assigns the corresponding hop_ID to all the sensor nodes. In Figure 3.6 the three nodes coming within the transmission range of BS are 2, 3 and 4 which are CH1, CH2, and CH3, respectively in the figure. The CHS_msg consists of the following parameters:

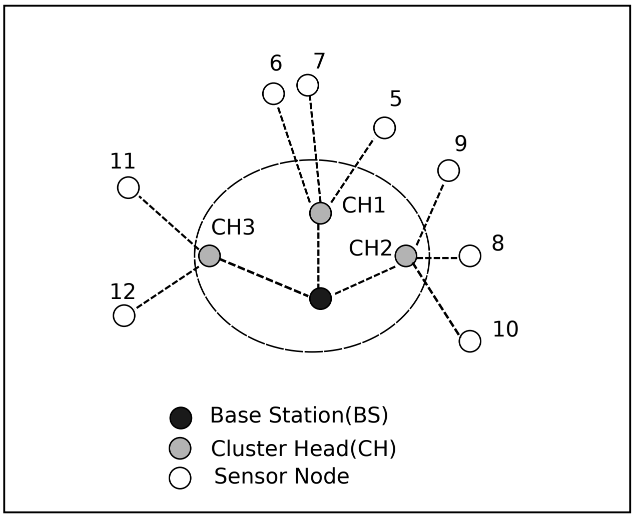

Figure 3.6.: Node Connected to BS within the R_t

1. hop_ID, the number of hops from the current node to the BS.
2. msg_type, the type of message send by the BS to the sensor network. It can be a CH selection message or it can be a query processing message.
3. ch_ID, the id of the previous CH or the BS.

Step 4: In Figure 3.6, CH 1 only send the CHS_msg to 5, 6 and 7 so in this case CH1 become CH.

Step 5: Once all the CHs are formed in the given network, on the basis of number of nodes connected to it, the number of hop required for every CH to reach the BS is calculated.

3.4.2. Data Transmission from Sensor to Sink

Whenever there is a need to transmit sensor data the reactive part of the protocol transfers the data from the sensor node to the *BS*. The proactive part of the protocol deals with transmission of data from the sensor to the*BS* at regular time intervals. The sensor nodes send their unique node identification number (n_i) and data to the CH assigned to them. The CHs aggregate data and send the information to the *BS* using minimum number of hops as designated by the hop_id(H_{id}) of the CH.

The *BS* would use the above information to process any type of application. In this task it may generate spatial and temporal data, process the desired parameters and predict events.

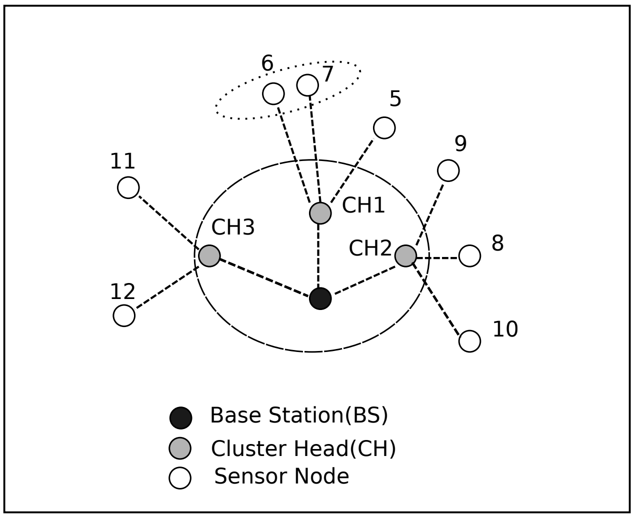

Figure 3.7.: All node connected to BS

3.4.3. Merging

If two nodes are sensing the same area or if the node is falling in the sensing range (R_s) of another node connected to the same CH then it is of no use to sense same data using multiple sensors. This reduces the energy efficiency of the network and as we have limited available operating power for wireless networks, this will be a major reason for shortening the lifetime of network. An improvement over this approach is the proposed merging algorithm. If nodes are falling in the same sensing range then these nodes should work like one node with higher initial energy as shown in Figure 3.7. Physically by merging we mean that these nodes will work in sleep and active mode with initial energy (E_0) higher than other nodes $(say 1.5J)$. This will save considerable amount of energy and hence lifetime of network increases. After the merging the above steps 3.3.1 to 3.3.3 are repeated again to set the network again.

3.4.4. Simulation and Results

The network is simulated in MATLAB. The parameters on which the efficiency of the network is compared or determined, is the energy required for trans receiving and processing in the network and the life time of the network. The energy of every node is evaluated using First Order Radio Model in Section 2.8.

Once the energy of the node required in transmitting and receiving is determined the number of round the node will last can be predicted. The number rounds till which the CHs are going to last determine the life time of the whole network.

We have simulated network using concept of merging. The simulations were performed for the rounds for First node to die-out and the Network to die-out. The results are shown in Table 3.3 and 3.4.It can be observed that when merging is used in the network, the number

Table 3.3.: Simulation Result

No. of Nodes	First node die-out (Rounds)		
	Original Network	After Merging	% Change
50	57	69	21 %
100	47	61	29.8 %
200	39	49	25.6 %

Table 3.4.: Simulation Result

No. of Nodes	Network die-out (Rounds)		
	Original Network	After Merging	% Change
50	74	83	12.1 %
100	91	101	11 %
200	105	120	14.3 %

of rounds for the First Node to Die in the network increases by 21 %, 29.8 % and 25.6% for network with 50, 100, and 200 nodes respectively. The network die-out time also increases in all the cases.

3.5. Summary

This chapter discusses an approach of sensing multiple attributes in an environment and then producing an alarm if any unexpected condition is predicted. The protocol used is hybrid and bidirectional. The dynamic CH formation increases the network life. Efficient data aggregation techniques are used so that the data transmission is energy efficient, with less aggregation overhead and is reliable. The use of data aggregation has minimized the data transmissions and also made it fast. The concept of distributed computing is also applied so that the work of the sink node is made easy and fast. The parallel processing in the application processor has increased the speed of the data analysis.

The modification of the above algorithm was proposed in the form of merging method which minimizes the global energy usage by making sensing nodes to work in sleep and active mode. Merging outperforms the conventional network's energy usage by uniformly distributing the load of sensing in the nodes falling in the same sensing range. Clearly our simulations show that

1. Network using merging is more energy efficient than the conventional network.

2. Lifetime of network using merging is increased.

The application of this type of network become limited in sparse deployment as the sensing range of the nodes is less. There has to be large number of nodes for feasibility of this network. Based on our MATLAB simulations we are confident that our proposed model will outperform the conventional models. Providing WSN's with such efficient models will open up a whole new horizon for them.

4. Hierarchical Agglomerative Clustering

4.1. Introduction

Hierarchical clustering algorithms create sets of clusters. They differ in how the sets are created. A tree data structure called dendrogram is used to illustrate the hierarchical clustering technique and the sets of different clusters. The space complexity of hierarchical algorithms is $O(n^2)$ as this is the space required for the adjacency matrix. The time complexity is $O(kn^2)$ because there is one iteration for each level in the dendrogram. HAC start with each individual item in its own cluster and iteratively merge clusters until all items belong to one cluster. The basic process of HAC is the merging of clusters based on their proximity [143].

The main advantages of the HAC approach are as follows [144]:

1. Simple computation and easy implementation.

2. Less restricted assumptions and more flexibility: HAC could use simple qualitative connectivity information of a network or quantitative data through Received Signal Strength (RSS) or GPS. In addition, other factors could easily be incorporated into the algorithm. For instance, different weights could be assigned to different nodes or connections for specific scenarios.

3. Less resource for clusters establishment: Using the HAC approach, nodes can finish the CHs election and announcement, cluster establishment, and scheduling at the same time. It can greatly reduce resource dissipation.

4. Without the need of periodic re-clustering or network updating: The HAC approach generates a logical CH backup chain during the cluster generation process. It makes clusters easily adaptive to network changes without extra information exchanges or the need of periodic announcement, such as CH.

Different hierarchical clustering techniques are classified as shown in the Figure 4.1.

Hierarchical algorithms are categorized as agglomerative or divisive [145]. Agglomerative implies that the clusters are created in a bottom-up fashion, while divisive algorithms work in

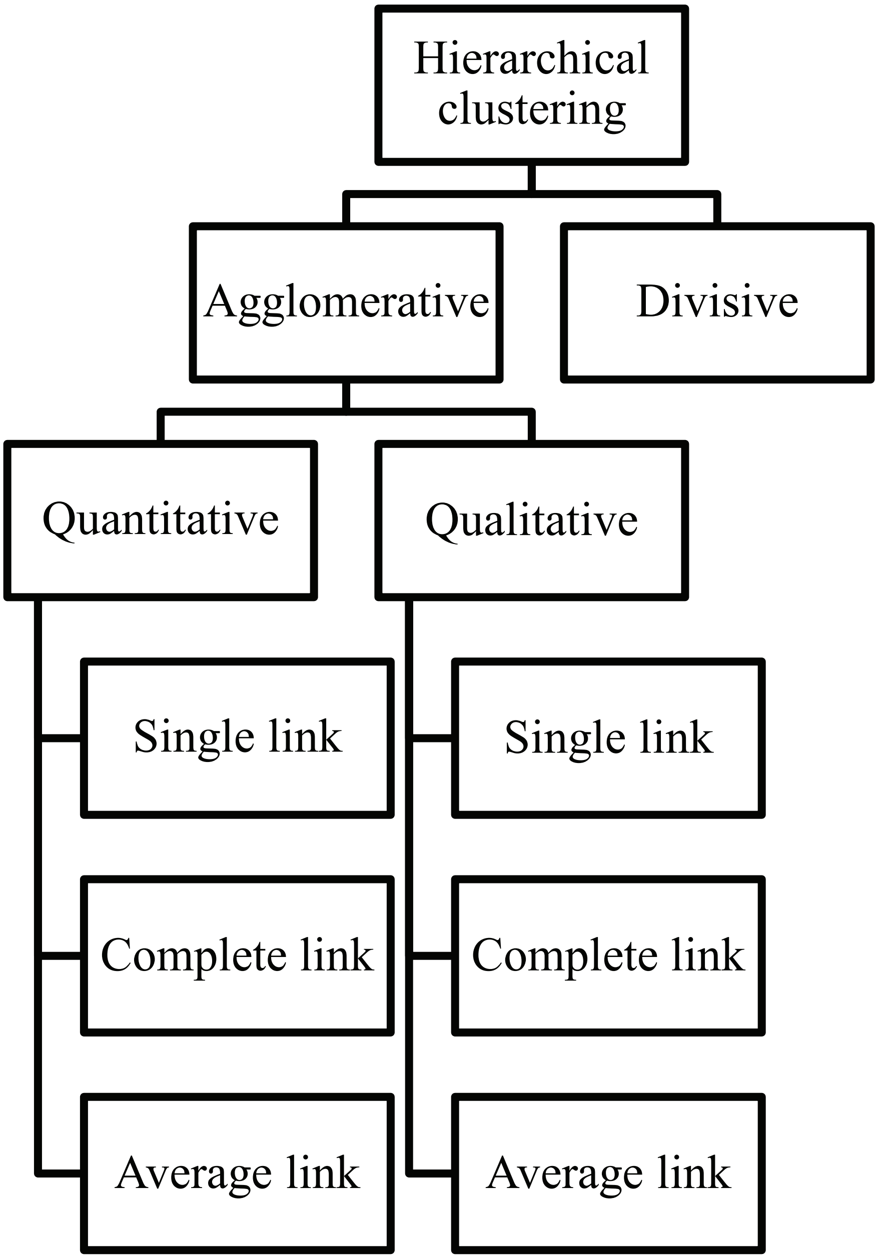

Figure 4.1.: Hierarchical clustering

a top-down fashion. Another descriptive tag indicates whether each individual sensor note is handled one by one, serial (sometimes called incremental), or whether all items are examined together, simultaneous.

The current chapter proposes a HAC algorithm to have minimum energy consumption in a WSN. Section 4.2 explains the Hierarchical Quantitative and Qualitative Clustering. Section 4.3 deals with HAC process and the section 4.4 talks about the simulation and results for qualitative and quantitative data.

4.2. Hierarchical Quantitative and Qualitative Clustering

The motivation of the research is to provide efficient clustering without requiring the global knowledge of network by using the bottom-up called the Hierarchical agglomerative clustering (HAC)[144]. With the bottom-up approach, sensing nodes build clusters before they select CHs. In this manner, the bottom-up approach can be a better way to implement self-organization, scalability and flexibility. If the distance between the sensing nodes is calculated using their location then it's quantitative HAC. If the received signal strength is used to calculate the distance between the nodes then it's qualitative HAC. This chapter compares the various agglomerative clustering techniques applied in a WSN. The simulations are done in MATLAB and the comparisons are made between the different protocols using dendrograms[14].

4.3. HAC Process

The process of HAC comprises three common key steps: obtaining the input data set, computation of the resemblance coefficients, and executing the clustering method.

4.3.1. Input Data Set

An input data set for HAC is a component-attribute data matrix. Components are the sensing nodes that are to be clustered based on their similarities. Attributes are the properties of the components. The attributes are the x and y coordinates of the sensing nodes that give the exact position of the node. The components or the attributes can be easily added or removed from the data set for different applications. Obviously, the more factors are considered, the more restricted assumptions and computations are needed. The type of input data set can be classified into quantitative data and qualitative data. The location information is used as the quantitative input data and the connectivity information is used as the qualitative data.

4.3.2. Computation of Resemblance Coefficients

A resemblance coefficient for a given pair of components indicates the degree of similarity or dissimilarity between these two components, depending on the way in which the data is represented. It could be quantitative or qualitative. For quantitative data, the current research work uses Euclidean distance as the similarity measure. Less the Euclidean distance and more similar the two sensing nodes and thus belong to the same cluster.

if $S_1 = (x_1, y_1)$ and $S_2 = (x_2, y_2)$ are the two nodes, then the distance (Dis) between $S_1$and S_2 is given by Eqn. 4.1.

$$Dis(S_1, S_2) = D_{1,2} = \sqrt{[(x_1 - x_2)^2 + (y_1 - y_2)^2]} \tag{4.1}$$

The distance between each pair of nodes is calculated and the adjacency matrix is obtained as given in Table 4.2. For quantitative data [146], following ways are used to calculate the resemblance coefficients Eqns.4.2-4.5:

- Jaccard:

$$SIM_{(a,b)} = M_{1-1}/(M_{1-1} + M_{1-0} + M_{0-1}) \tag{4.2}$$

- Sorenson:

$$SIM_{(a,b)} = 2M_{1-1}/(2M_{1-1} + M_{1-0} + M_{0-1}) \tag{4.3}$$

- Simple Matching:

$$SIM_{(a,b)} = M_{1-1} + M_{1-0}/(M_{1-1} + M_{1-0} + M_{0-1} + M_{0-0}) \tag{4.4}$$

- Dice:

$$SIM_{(a,b)} = M_{1-1}/(M_{1-1} + 0.5M_{1-0} + 0.5M_{0-1}) \tag{4.5}$$

where M is the attribute which comes from the entries in Connectivity matrix based on the direct connection of the node,

M_{1-1}: represents the total number of attributes where a and b both have a value of 1,

M_{0-0}: represents the total number of attributes where a and b both have a value of 0,

M_{0-1}: represents the total number of attributes where the attribute of a is 0 and the attribute of b is 1,

M_{1-0} : represents the total number of attributes where the attribute of a is 1 and the attribute of b is 0.

Dissimilarity coefficient

$$DSIM_{(a,b)} = 1 - SIM_{(a,b)} \tag{4.6}$$

In the simulation, the clustering of the qualitative and quantitative data is done using single link, complete link and average link methods as explained in the next sections.

4.3.3. Execution of the HAC Method

Execution of HAC involves various steps and each step merges two clusters together and updates the Resemblance Matrix. Updating the Resemblance Matrix is an important step and various methods could be adopted. With the same data set, we may get different clustering results by using different HAC algorithms. The type of algorithm depends on how the distance between the motes in two clusters is calculated. This is not an easy task as there are many interpretations for the distance between clusters. There are three main types of HAC [8]:

1. Single link: It's also called as nearest neighbor method. Two clusters are merged if the minimum distance between any two points is less than or equal to the threshold distance being considered, It considers the smallest distance between a node in one cluster and a node in the other. Thus

   $$Dis(K_i, K_j) = min(Dis(S_{il}, S_{jm})) \bigvee S_{il} \in K_i \notin K_j$$

   and

   $$S_{jm} \in K_j \notin K_i \tag{4.7}$$

   where $K_i$and K_j are two different clusters in the sensor network N, $S_{pq}$is the q^{th} sensor node of p^{th} cluster.

2. Complete link: Although it's similar to the single link algorithm, the difference lies in the fact that it considers the largest distance between a node in one cluster and a node in the other. Thus

   $$Dis(K_i, K_j) = max(Dis(S_{il}, S_{jm})) \bigvee S_{il} \in K_i \notin K_j$$

   and

   $$S_{jm} \in K_j \notin K_i \tag{4.8}$$

3. Average link: The average link technique merges two clusters if the average distance between any two points in the two target clusters is below the distance threshold. It takes the average distance between a node in one cluster and a node in the other. Thus

   $$Dis(K_i, K_j) = mean(Dis(S_{il}, S_{jm})) \bigvee S_{il} \in K_i \notin K_j$$

and

$$S_{jm} \in K_j \notin K_i \tag{4.9}$$

4.3.4. Single Complete and Average Link Clustering Algorithms

The steps involved in single link clustering algorithms are as follows:

1. Find the minimum distance in the adjacency matrix.
2. Cluster the two nodes with the minimum distance.
3. The distance between the clustered nodes is calculated with the rest of the unclustered nodes.
4. Let's say i and j nodes are clustered since $D_{(i,j)}$ was minimum in the adjacency matrix.
5. Now the adjacency matrix is updated as
   a) $D_{(i,k)}$and $D_{(j,k)}$ is replaced with $min(D_{(i,k)}, D_{(j,k)}) \forall k \in N$
   b) The dimension of the adjacency matrix is also reduced by one.
6. The steps 1-6 are repeated until the adjacency matrix is left with 2*2 elements.

The complete link algorithm works in the same way except the 5^{th} step is replaced with $max(D_{(i,k)}, D_{(j,k)}) \forall k \in N$. The dimension of the adjacency matrix is also reduced by one. In the average link algorithm the distance is the average of the two distances and replaced with $mean(D_{(i,k)}, D_{(j,k)}) \forall k \in N$.

4.4. Simulation and Results

Some of the assumptions made during simulation are as follows:

1. The nodes in the network are quasi-stationary, i.e. location of neighbour need not be determined every time clustering is performed.[10]
2. Propagation channel is symmetric.
3. Nodes are left unattended after deployment.
4. All nodes have similar capabilities, processing, communication and initial energy.
5. The coordinates of the nodes deployed is known.
6. The nodes are deployed at fixed locations.

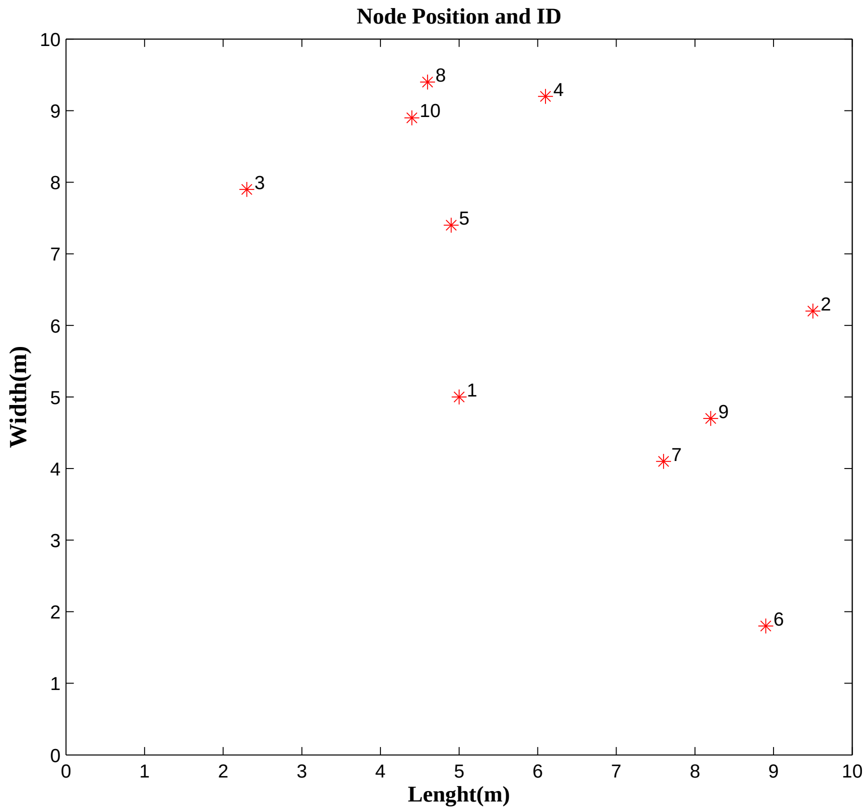

Figure 4.2.: plot of the randomly deployed sensor nodes

Table 4.1.: X and Y coordinate of the nodes

Node ID	1	2	3	4	5	6	7	8	9	10
X data	5	9.5	2.3	6.1	4.9	8.9	7.6	4.6	8.2	4.4
Y data	5	6.2	7.9	9.2	7.4	1.8	4.1	9.4	4.7	8.9

7. For depiction purpose 10 nodes are used.

Ten sensing nodes were randomly deployed in a room. The sensing nodes are clustered based on hierarchical agglomerative qualitative and quantitative clustering. As mentioned earlier for quantitative data, Euclidian method is used to calculate the resemblance coefficients and for qualitative data, Sorensen method is used to calculate the resemblance coefficients. The plot of the deployed nodes is depicted in the Figure 4.2.

4.4.1. Quantitative Data

Firstly for quantitative data, the adjacency matrix is calculated, based on Euclidian distances. The x and y coordinates of the sensors are depicted in Table 4.1. The adjacency matrix is as shown in the Table 4.2.

Table 4.2.: Adjacency matrix for quantitative data

Node ID	1	2	3	4	5	6	7	8	9	10
1	0	4.66	3.96	4.34	2.40	5.04	2.75	4.42	3.21	3.95
2	-	0	7.40	4.53	4.75	4.44	2.83	5.86	1.98	5.77
3	-	-	0	4.01	2.64	8.99	6.52	2.75	6.71	2.33
4	-	-	-	0	2.16	7.90	5.31	1.50	4.97	1.73
5	-	-	-	-	0	6.89	4.26	2.02	4.26	1.58
6	-	-	-	-	-	0	2.64	8.73	2.98	8.40
7	-	-	-	-	-	-	0	6.09	0.85	5.77
8	-	-	-	-	-	-	-	0	5.92	0.54
9	-	-	-	-	-	-	-	-	0	5.66
10	-	-	-	-	-	-	-	-	-	0

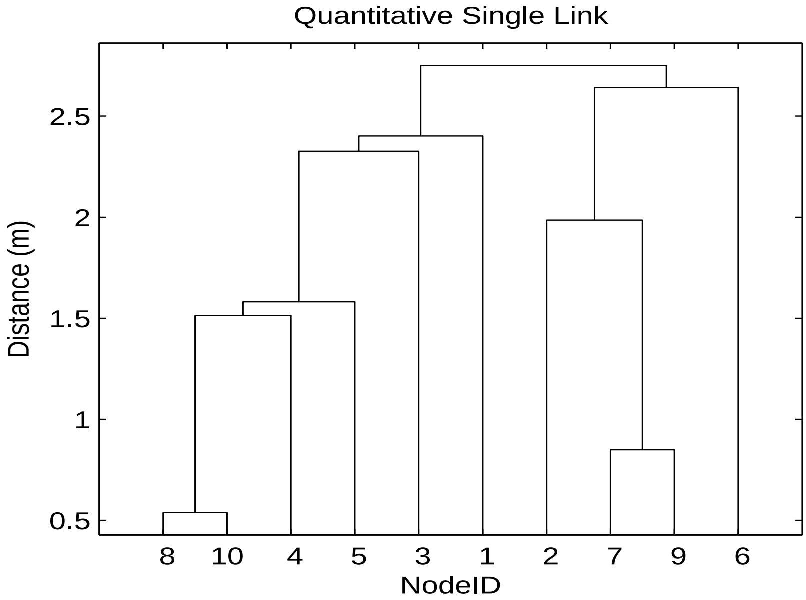

Figure 4.3.: Single link HAC for quantitative data

Based on the above distance matrix, clustering is done using the three approaches of single link, complete link and average link. Correspondingly the Figures 4.3, 4.4 and 4.5 represent the dendrograms formed in single link, complete link and average link respectively.

Biggest disadvantage in the single link HAC is that the clustering creates clusters with long chain as shown in the Figure 4.3.

4.4.2. Qualitative Data

In the case of qualitative data, the x and y coordinates are not required to be calculated. The current research work uses the one-hop network connectivity data as the qualitative input data, where the “1” value represents a one-hop connection and the “0” value represents

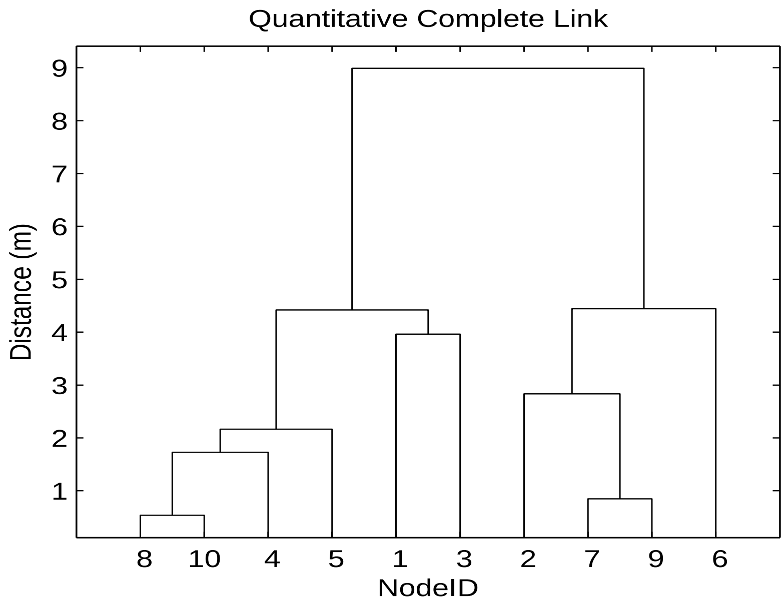

Figure 4.4.: Complete link HAC for quantitative data

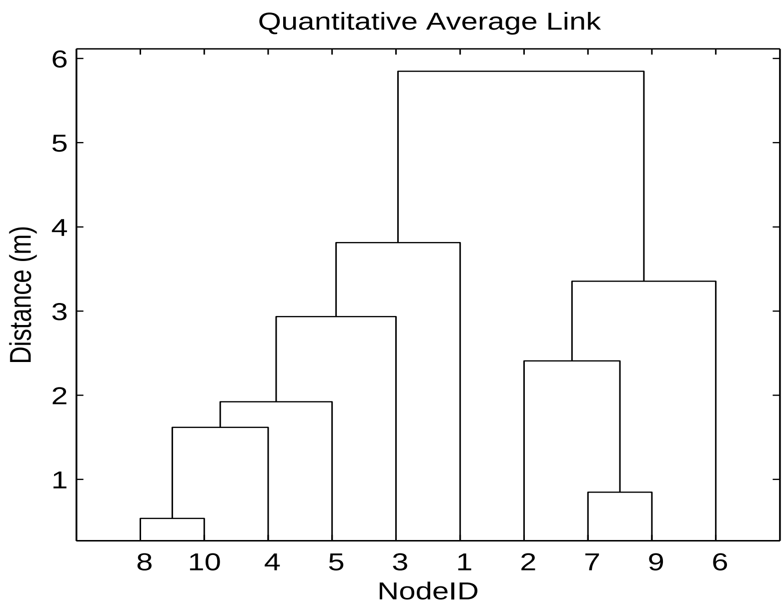

Figure 4.5.: Average link HAC for quantitative data

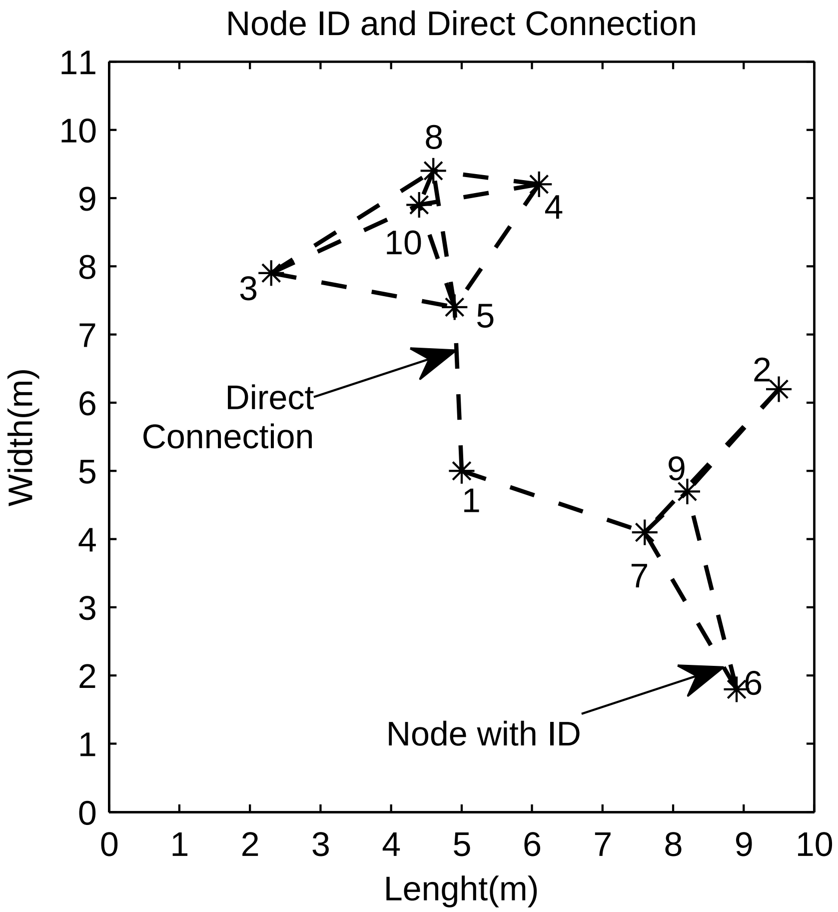

Figure 4.6.: One hop connectivity diagram

no direct connection. The received signal strength (RSS) is used to find out the network connectivity. If two nodes are within the transmission range then it's a one hop connection. The Table 4.3 is the connectivity matrix obtained by the one hop connectivity of the nodes is shown in the Figure 4.6.

The adjacency matrix obtained by Sorenson coefficients is shown in the Table 4.4.

Based on the above distance matrix, clustering is done using the three approaches of single link, complete link and average link. Correspondingly the Figures 4.7, 4.8 and 4.9 represent the dendrograms formed in single link, complete link and average link respectively for the quantitative data.

Table 4.3.: Connectivity matrix

Node ID	1	2	3	4	5	6	7	8	9	10
1	1	0	0	0	1	0	1	0	0	0
2	0	1	0	0	0	0	1	0	1	0
3	0	0	1	0	1	0	0	1	0	1
4	0	0	0	1	1	0	0	1	0	1
5	1	0	1	1	1	0	0	1	0	1
6	0	0	0	0	0	1	1	0	1	0
7	1	1	0	0	0	1	1	0	1	0
8	0	0	1	1	1	0	0	1	0	1
9	0	1	0	0	0	1	1	0	1	0
10	0	0	1	1	1	0	0	1	0	1

Table 4.4.: Adjacency matrix by Sorenson dissimilarity coefficients

Node ID	1	2	3	4	5	6	7	8	9	10
1	0	0.56	0.33	0.45	0.14	0.56	0.33	0.45	0.33	0.33
2	-	0	1	1	0.64	0.33	0.33	1	0.33	1
3	-	-	0	0.09	0.14	1	0.67	0.09	0.67	0
4	-	-	-	0	0.23	1	0.82	0	0.82	0.09
5	-	-	-	-	0	0.64	0.43	0.23	0.43	0.14
6	-	-	-	-	-	0	0.33	1	0.33	1
7	-	-	-	-	-	-	0	0.82	0	0.67
8	-	-	-	-	-	-	-	0	0.82	0.09
9	-	-	-	-	-	-	-	-	0	0.67
10	-	-	-	-	-	-	-	-	-	0

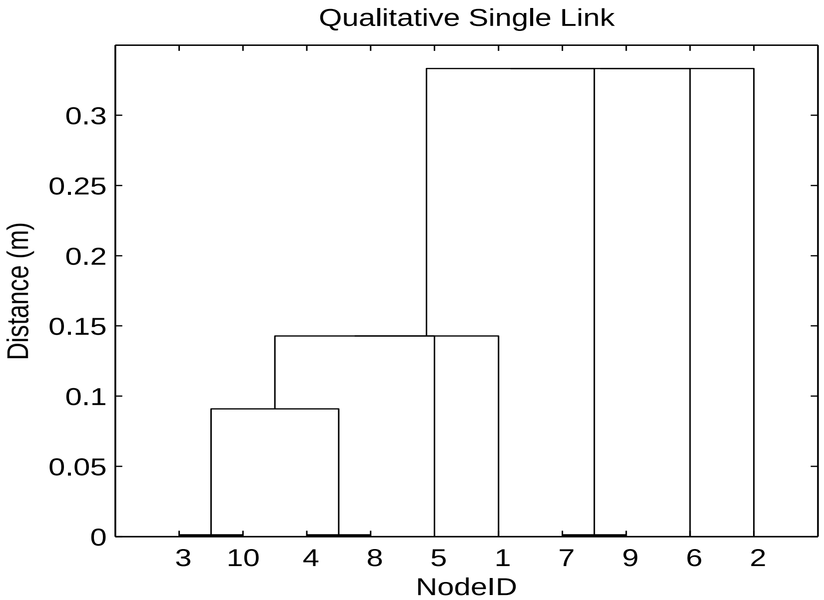

Figure 4.7.: Single link HAC for qualitative data

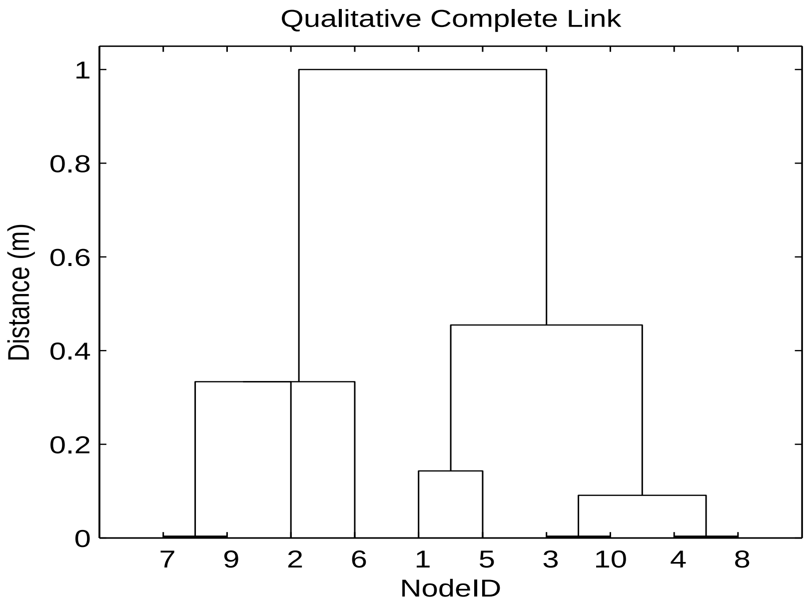

Figure 4.8.: Complete link HAC for qualitative data

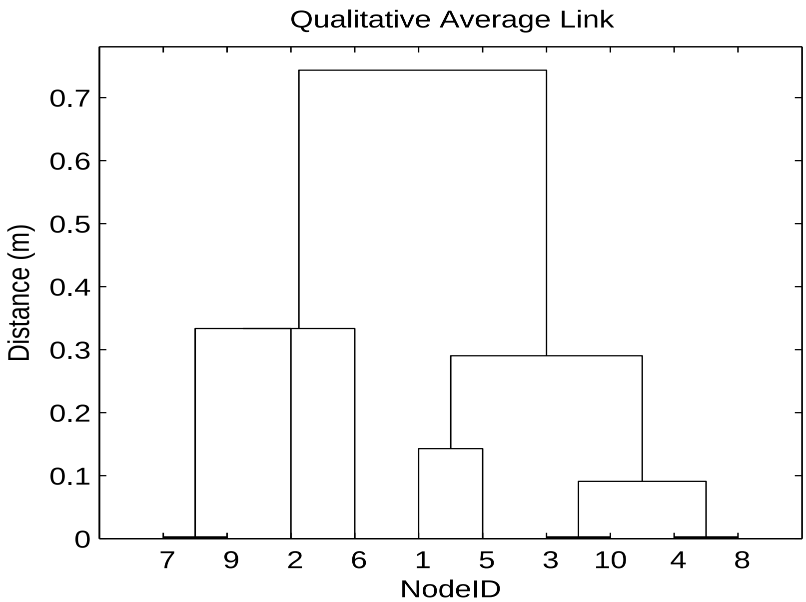

Figure 4.9.: Average link HAC for qualitative data

4.5. Results and Observations

A wireless sensor network consists of a large number of sensing nodes randomly deployed in a region. Clustering not only reduces energy consumption but also achieves better network performance. Hierarchical agglomerative clustering uses bottom up approach and hence the complete network information is not required. The dendrograms for the various HAC protocols for qualitative and quantitative data were shown in the previous section. . If the position of the nodes is not known due to the issues like cost, time, energy conservation then the Sorensen method is used to find out the similarity matrix and then perform the clustering of the nodes. The dendrograms for single and complete link protocol clearly shows that single link has a chain effect, making the clustering inefficient. The dendrograms of quantitative single link, complete link, and average link are almost similar to the dendrograms of qualitative single link, complete link and average link respectively. So, it shows that there is no need to find out the location of the nodes in a randomly deployed WSN. Finding out the position of the nodes not only make the clustering complex but also consumes a lot of energy. When the number of nodes increase or behave asymptotically, then it becomes computationally complex to calculate the adjacency matrix. Hence in such cases, Qualitative approach is used.

4.6. Summary

Clustering not only reduces energy consumption but also achieves better network performance. Agglomerative hierarchical clustering is a bottom-up clustering method where clusters have sub-clusters, which in turn have sub-clusters, etc. The dendrograms of quantitative single link, complete link, and average link are almost similar to the dendrograms of qualitative single link, complete link and average link respectively. So, it demonstrates that there is no need to find out the location of the nodes in a randomly deployed WSN. Finding out the position of the nodes not only make the clustering complex but also consumes a lot of energy.

5. Ensuring Coverage and Connectivity in Single Sink

5.1. Introduction

Wireless Sensor Network (WSN) is a network of densely deployed large number of sensor nodes. WSNs are basically data gathering networks in which data are highly correlated and the end user needs a high level description of the environment sensed by the nodes [1]. WSNs are deployed to monitor physical events or the state of physical objects such as bridges in order to support appropriate reaction to avoid potential damages [2]. The nodes and the related protocols in a WSN should be designed to be extremely energy efficient as battery recharging may be impossible[51]. In direct communication WSN, the sensor nodes directly transmit their sensed data to the Base Station (BS) or sink without any coordination between the two. However, in cluster based WSNs, the network is divided into clusters. Each node exchanges its information only with its cluster head (CH) which transmits the aggregated information to BS. Data aggregation at CHs causes a significant reduction in the amount of data sent to the BS and results in saving both energy and bandwidth resources [10].

The most important phase of cluster-based routing protocols is the CH selection procedure that ensures uniform distribution of energy among the sensors, and consequently increasing the lifespan of a sensor network [12]. Once the CH are identified, they form a backbone network to periodically collect, aggregate, and forward data to the BS using the minimum energy (cost) routing. This method significantly enhance the network lifetime compared to other known methods. The major challenges include equal distribution of each cluster over the entire sensor network and the energy dissipation caused by the frequent information exchange between selected CH and nodes in the cluster in every setup phase of cluster formation [11, 4].

If CH is selected on the basis of the concept of maximum number of nodes connected, then it may happen that one or more unique nodes are not connected to any of the selected CHs. In such case some outliers may be created which are not connected to any of the CHs, although they are in the transmission range. The proposed algorithm [13]deals with the CH selection based on the unique node concept. A unique node is the one which is not connected to any

Table 5.1.: Notations used

Notation	Meaning	Notation	Meaning
R_s	Sensing Range	R_t	Transmitting Range
n_0	Node id of BS	n_i	Node id of ith node
E_{DA}	Data aggregation energy	E_r	Residual Energy < E_0
E_0	Initial node energy (1J)	$E_{sensing}$	Sensing energy
k	Number of bits in one packet	N	Total nodes in the network
E_{th}	Energy threshold value at which the CH selection restart		
Nb_i	ith hop neighbour of BS or neighbour of nodes in (i-1)th hop		
d_{ij}	Distance between ith node to jth node		
N_i	Set of neighbouring nodes of ith node		
CH_{id}	Cluster head ID's for each nodes		
$Count_i$	The number of neighbours of ith node		

other CHs. The algorithm proposed uses two parameters, namely, number of neighbouring nodes and the residual energy for the selection of CH in WSN. The details are mentioned in the subsequent sections.

The section 5.2 gives the system model and the assumption made. The section 5.3 deals with the proposed clustering algorithm. The subsequent sections 5.4 and 5.5 deal with the simulation and results and the summary of the chapter respectively.

5.2. System Model and Assumptions

The notations used are detailed in the Table 5.1.

The network considered has the following properties:

1. The nodes are homogenous with initial energy of 1Joules
2. The nodes transmit the data to the BS in multiple hops.
3. The hops are determined based upon distance from the BS.
4. The sensors used have transmitting range of 100-150m (Outdoor) and 50-75m (Indoor). If R_t (150 m) and Rs (75 m) represents transmission and sensing range respectively, it is assumed that $R_t \geq 2R_s$ as given in [12]. If there is one more sensor within the transmission range of a sensor, then the sensed information is same for both the sensors as shown in Figure 5.1.
5. The nodes are randomly deployed.
6. Sensor nodes are stationary unlike mobile nodes architecture proposed in the research paper [147].

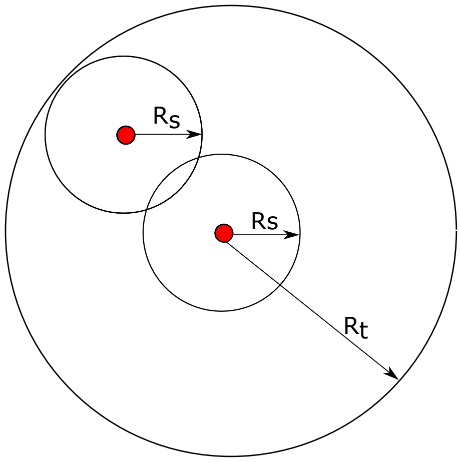

Figure 5.1.: Sensing and Transmission range

7. BS (n_0) is fixed and installed somewhere in the middle of the network (250,250) unlike research paper which deals with *BS* identification [148].

8. The nodes are proactive, i.e., they transfer data to the BS in periodic intervals.

9. The CHs are selected initially at the onset of the network. The CH is re-selected or rotated as per the proposed algorithm if the energy of the CH falls below the threshold.

10. The energy of every node is evaluated using First Order Radio Model in Section 2.8.

5.3. Proposed Clustering Algorithm

On the basis of location of various nodes, the proposed algorithm identifies the clusters and CHs. The CHs are chosen on the basis of unique neighbour nodes and their residual energy. After CH identification, the data transmission from a specific node is done using CHs till it reaches BS. Each CH combines the data collected from its connected nodes and performs data aggregation that reduces the amount of data to be transmitted. The aggregated data is sent to the BS through the intermediate CHs as per routing table established earlier. The stepwise algorithm is described in the following section.

Step 1: N number of motes is randomly deployed in the region to be sensed. The BS is given the id of 0 and it is manually located on the network field.

$N = \{n_1, n_2, n_3, ...n_N\}$

Step 2: Calculate the set of neighbour nodes Nb_i and the number of neighbour nodes $count_i$ for the i^{th} node n_i on the basis of the transmission range as depicted in the Table 5.2.

The function neighbour_info gives the result as shown in the Figure 5.2.

Step 3: BS always behaves as the CH. CH selection (CHS) process starts from BS. BS transmits the CH Selection message (*CHS_msg*)to all the neighbouring nodes. Nodes having

Algorithm 5.1 Neighbour_info

$function[count_i, Nb_i] = neighbour\_info(n_i);$

The function $neighbour\_info$ returns all the neighbouring nodes of a node n_i The distance between a given sensor n_i and n_j is given by

$d_{ij} = \sqrt{}[(x_i - x_j)^2 + (y_i - y_j)^2]; i = 1, 2, \ldots, N; j = 1, 2, \ldots, N;$

$\{Nb_i | d_{ij} < R_t \forall j, j \in N\}$

$Nbi = \{n_1, n_2, n_3, \ldots n_n\};$

set of neighbour nodes IDs within the transmitting range (R_t). $neighbor\_info$ function is repeated for all N nodes and the BS.

Table 5.2.: Neighbour Info Details for first hop

i	n_i	$count_i$	Nb_i
0	n_0	3	n_1, n_2, n_3
1	n_1	3	n_0, n_2, n_7
2	n_2	6	$n_0, n_1, n_4, n_5, n_7, n_{10}$
3	n_3	5	$n_0, n_{11}, n_{12}, n_{13}, n_{14}$

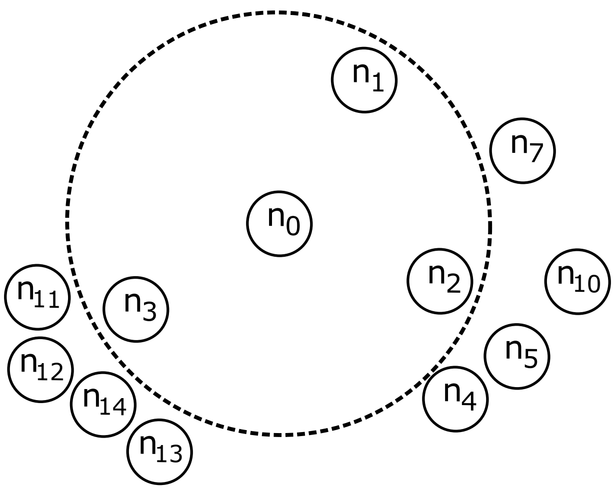

Figure 5.2.: Step 3, Neighbours

Algorithm 5.2 All_neighbours

$Function\ [Neighbours] = all\_neighbours(Nb_i) \bigcup Nb_i$
$Neighbours\, for Nb_0 = \{n_1 n_2 n_3\} \bigcup \{n_0 n_2 n_7\} \bigcup \{n_0 n_1 n_4 n_5 n_7 n_{10}\} \bigcup \{n_0 n_{11} n_{12} n_{13} n_{14}\}$
$= \{n_1 n_2 n_3 n_4 n_5 n_7 n_{10} n_{11} n_{12} n_{13} n_{14}\}\ except\, n_0$

residual energy more than the threshold energy are eligible to become a CH. The step 3 is depicted in the Figure 5.2. All_neighbours function returns the set of all the neighbouring nodes $[Neighbours]\, of\, Nbi$

Step 3_A: This step is applicable only when the BS determines its neighbour and set the BS position according to the below equation to minimize the BS to neighbour node distance to save the energy and reduce the synchronization time; $\sum_{i=1}^{count_i} d_i \to min$. where d_i is the distance from a node toi^{th} node. The optimal BS coordinates are then given by

$$(x_0, y_0) = arg\,min \underset{(x,y)}{min} \sum_{i=1}^{N} \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \tag{5.1}$$

The minimumis obtained by setting the partial derivatives to zero:

$$\frac{\partial}{\partial x} \sum_{i=1}^{count_i} d_i = \frac{\partial}{\partial x} \sum_{i=1}^{count_i} d_i = 0\ \ for\ \ x = x_0,\ y = y_0 \tag{5.2}$$

The partial derivatives are

$$\frac{\partial}{\partial x} \sum_{i=1}^{count_i} d_i = \sum_{i=1}^{count_i} \frac{\partial}{\partial x} d_i = \sum_{i=1}^{count_i} \frac{x_i - x}{d_i} \tag{5.3}$$

$$\frac{\partial}{\partial y} \sum_{i=1}^{count_i} d_i = \sum_{i=1}^{count_i} \frac{\partial}{\partial y} d_i = \sum_{i=1}^{count_i} \frac{y_i - y}{d_i} \tag{5.4}$$

By using the vector notations, the vector pointing to the location of the i^{th} sensor node is $n_i = (x_i, y_i)$, and the distance vector between the sensor node and sink is $d_i = n_i - s = ((x_i, x), (y_i, y))$. Let e_i be the unit vector from i^{th}sensor node towards the nearest sink (the orientation vector); that is,

$$e_i = \frac{d_i}{d} = \frac{1}{d}((x_i, x), (y_i, y)). \tag{5.5}$$

where $d = |d_i|$ to normalize the magnitude of e_i. Using the above three equations $r = \sum_{i=1}^{count_i} e_i = 0$ for $s = (x_0, y_0)$, that is, the average distance is minimized if the resultant r of the orientation vector is zero,

In Neighbours the CH_{id} or (initially BS ID) is excluded.

Step 4: The function $unique\_neighbor\_info$ returns the unique nodes connected to a node.

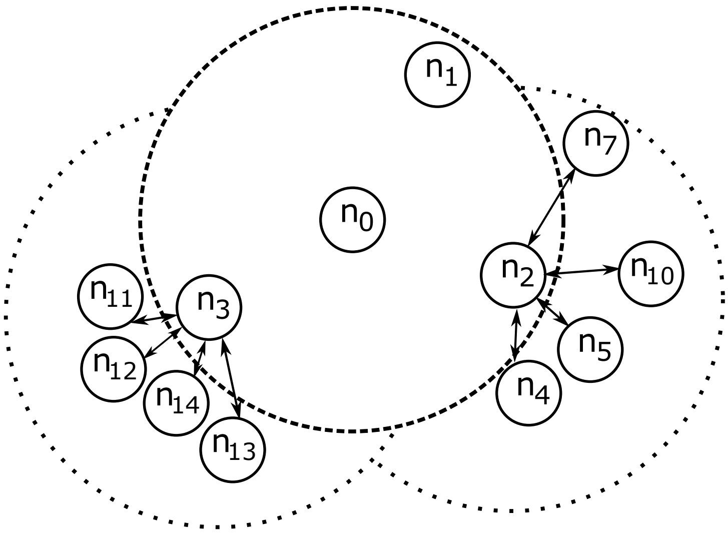

Figure 5.3.: Step 4, Nub_i

Table 5.3.: Unique neighbour details for first hop

i	n_i	$count_i$	Nb_i	$ucount_i$	Nub_i
1	n_1	3	$n_0\, n_2\, n_7$	0	-
2	n_2	6	$n_0\, n_1\, n_4\, n_5\, n_7\, n_{10}$	3	$n_4\, n_5\, n_{10}$
3	n_3	5	$n_0\, n_{11}\, n_{12}\, n_{13}\, n_{14}$	4	$n_{11}\, n_{12}\, n_{13}\, n_{14}$

The unique nodes are the ones which are not connected to any other nodes in $(i+1)^{th}$ hop as shown in the Figure 5.3.

Table 5.3 gives the details of the unique neighbours for first hop.

A flag is set for all the Nb_i nodes appearing in the set Neighbours for Nb_i

Step 5: Consider the unflagged elements of Neighbours. The function $set\_of\_neighbours$ returns NNb_i which is the set of all the neighbours of members of Nb_i depicted in Table 5.4.

The $Left\_neighbours$ contains the nodes whose cluster ID flag is not set by unique count.

$min_setflag_count$is a function which calculates the minimum of $count_i$of all the nodes in $Nb_i$for which the flag is not set.

Step 6: The steps 3 to 5 are repeated till all the nodes are covered by elected CH. All nodes should have a path up to BS through single or multiple hops depending upon the distance of the node from the BS.

Table 5.4.: Neighbour Info Details for first hop.

NNb_i	n_i		Nb_i
	n_0	NNb_0	$n_1 n_2 n_3$
NNb_0		NNb_1	$n_0 n_2 n_7$
NNb_0		NNb_2	$n_0 n_1 n_4 n_5 n_7 n_{10}$
NNb_0		NNb_3	$n_0 n_{11} n_{12} n_{13} n_{14}$

Algorithm 5.3 Unique_neighbour_info

$Function[ucount_i, Nub_i] = unique\_neighbor\_info(Nbi, Neighbours);$
where$Nubi \in Nbi$
$for\, i = 0\, to\, Nb_i$
{
  $ucount_i > 0\ then$
  {
    $noden_i is the CH$
    $set\ flag\ for\ Nb_i in\ the\ set\ Neighbours(Nb_i)$
    $set\ CH_{id}\ (n_i)\ for\ Nb_i\ in\ the\ set\ Neighbours(Nb_i)$
  }
}

Algorithm 5.4 Set_of_neighbours

$Function[NNb_i] = set\_of\_neighbors(Nbi);$
$NNbi$ is depicted in Table [8]
In the Table $N(Nb_0)$ is
$N(Nb_i)\ for\ i = 0$
= $N(Nb_0)$
= $N(n_1\, n_2\, n_3)$
= Nb_1, Nb_2, Nb_3
=$\{n_0, n_2, n_7\}, \{n, n_1, n_4, n_5, n_7, n_{10}\}, \{n_0, n_{11}, n_{12}, n_{13}, n_{14}\}$

Algorithm 5.5 Notflag_neighbours

$Function[Left\_Neighbours] = notflag\_neighbors(Neighbours, Setflag);$
where$Nubi \in Nbi$
$for\, i = 0\, to\, Neighbours$
{
  $Min\_count = min\_setflag\_count(Nb_i)$
  $node\ n_i is\ the\ CH$
  $set\ flag\ for\ Nb_i in\ the\ set\ Neighbours(Nb_i)$
  $set\ CH_{id}\ (n_i)\ for\ Nb_i\ in\ the\ set\ Neighbours(Nb_i)$
  $notflag\_neighbours(Neighbours, Setflag)$
  set $Neighbours = left\_neighbours$
}

Step 7: After clustering, if in the process of transmission of data, one of the CHs die out, then the CH at the previous hop comes to know about it since the data from the died CH didn't reach it. Suppose the CH at n^{th} hop dies out, then in such a case the clustering algorithm is repeated after $(n-1)^{th}$hop for the entire network.

Step 8: If the residual energy of the CH becomes less than the threshold energy (E_{th}) then the CH selection process needs to be re-instantiated. Suppose the residual energy of the CH at nth hop is less than the threshold, then the clustering algorithm is repeated after $(n-1)^{th}$ hop for that particular path.

Step 9: Once the routing path is established the data transmits through multi-hop. Each CH combines the data collected from its connected nodes through data aggregation. Steps are represented by flowchart in Figure 5.4.

Energy conservation is of prime consideration in sensor network protocols in order to maximize the network's operational lifetime [149]. In WSNs, the communication cost is often several orders of magnitude larger than the computation cost [150]. To avoid too much of communication, the CH aggregates the received data and transmits only a small data to the BS. Data aggregation is any process in which information is gathered and expressed in a summary form. Many research papers[151, 40, 152] have shown that aggregation at the CH considerably reduces the amount of data routed through the network, increasing the throughput and extending the lifetime of the sensor networks. Data aggregation also solves the purpose of estimating a missing value from a sensor. Sometimes it may happen that data from one of the sensors didn't reach the CH, in such cases the Jackknife estimate can be used to predict the value of the sensed parameter. It also performs fault tolerance. If the data received from one of the sensors doesn't match the estimated value, then accordingly the correction is made. The flow chart of the complete proposed algorithm is shown in the Figure 5.4.

5.4. Simulation and Results

Simulations are carried out to evaluate the proposed algorithm in MATLAB. Simulation is done with nodes placed randomly using uniform distribution throughout the network of dimension 500 m × 500 m. The BS is located at x = 250 and y = 250. The simulation parameters considered are as described in Table 5.5. The transmission cluster radius is taken as 150 m and initial threshold energy $E_{th} = E_0/2$.

The proposed algorithm is evaluated using the following measures; network lifetime and connectivity. After some rounds of transmission, when the residual energy of all the nodes approaches to E_{th} then the network adaptively reduces the value of E_{th}, thus increasing the network lifetime. Data aggregation at the CHs further enhances the network lifetime by

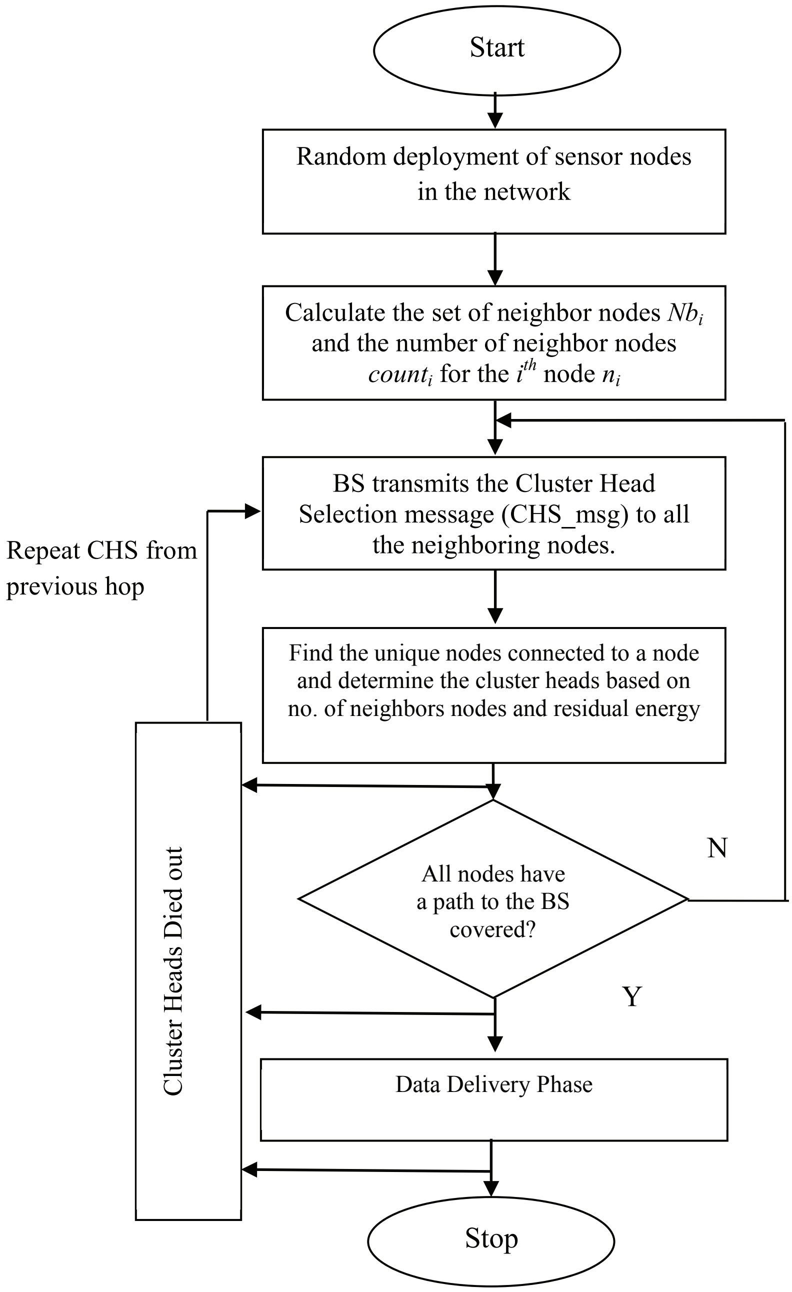

Figure 5.4.: Flow chart

Table 5.5.: Simulations Parameters values

Parameter	Value	Parameter	Value
N	$50,100$	R_s	$Point Sensor$
E_0	$Intial\, 1\, Joule$	R_t	$100m, 150m$
n_0	$(250,250)$	E_r	$E_0 - E_{proc} - E_{DA} - E_{sensing}$
n_i	$NodeID\, of\, i^{th}\, node$	E_{th}	$0.5J\, to\, 0.1J$
A	$(0,0)\, to\, (250,250)$		

Table 5.6.: Simulations Results and Comparison

Protocols	Initial Energy Joules	Rounds		
		First Node Died (FND)	10 % Nodes Died	90 % Nodes Died
DTE	1J	217	276	468
LEACH	1J	1848	2007	2570
Multi-Hop	1J	1936	2571	4112
Proposed	1J	2400	3000	6741

reducing the size of the data to be transmitted by the nodes. When the CH residual energy is less than equal to the threshold energy, only then the CH selection algorithm is carried out for n-1 hops. Thus, the proposed algorithm ensures that energy is not wasted in CH selection for every round. The network connectivity is effectively handled by the algorithm. The CHs connect all the nodes of the network. The proposed algorithm ensures that none of the unique nodes are left behind without being connected to the network. Hence the possibility of having outliers is nullified.

The proposed algorithm is compared with Direct Transmission Energy (DTE), LEACH and Multi – hop routing. For each protocol, nodes are randomly deployed by generating random coordinates using uniform distribution. For each protocol, 100 iterations were performed and the result is their average. We performed simulation on 50, and 100 nodes. The values of the simulation parameters are shown in the Table 5.5 and simulation results and comparison table is shown in the Table 5.6.

From Table 5.6, it is observed that the proposed protocol improves the network performance as compared to existing protocols. A huge improvement is observed for the rounds for first node to die, 10% nodes to die and 90% nodes to die as compared to DTE. However, if the results of proposed algorithm are compared with the best existing case, i. e. Multi-Hop case, it is observed that an improvement of 24%, 16.7% , and 63% is achieved for fisrt node to die, 10% nodes to die and 90% nodes to die respectively. Above results clearly show that the proposed algorithm gives better results as compared to existing methods. In this case the stability period is increased because the CH is rotating, only when the threshold value reaches 0.1 J.

5.5. Summary

The proposed algorithm for CH selection in a WSN using unique node concept has many advantages. The latency in transmitting the data in a single hop is much more than in the proposed multi hop WSN. Each node in the network transmits the data to CH/BS closest to it. The CH in turn transmits the data to the next CH, if required, to reach the BS. If the CH is selected on the basis of the concept of maximum number of nodes connected, then it may happen that one or more unique nodes are not connected to any of the selected CHs. Thus this algorithm deals with the CH selection based on the unique node concept. In the proposed algorithm there is no possibility of having any outlier, as all the unique nodes are connected to some or the other CH. Adaptability is well taken care of. After clustering, if in the process of transmission of data, one of the CHs die out, then the CH at the previous hop comes to know about it since the data from the died CH didn't reach it and the clustering algorithm is repeated after$(n-1)^{th}$ hop for the entire network. If the residual energy of the CH becomes less than the threshold energy then the CH selection process is re-instantiated after $(n-1)^{th}$ hop for that particular path. The Table 5.6 clearly depicts that the proposed clustering algorithm increases the network lifetime. In WSNs, the communication cost is often several orders of magnitude larger than the computation cost, thus the CHs perform data aggregation to reduce the amount of data to be transmitted.

Part II.

Multiple Sinks Wireless Sensor Network

6. Increasing Life Time in Multiple Sinks

6.1. Introduction

A noteworthy test in understanding a Wireless Sensor Network is energy proficiency. Since the sensing nodes are small in size, cost-effective, low-control gadgets, and have constrained battery power supply, energy utilization is an essential component while planning a WSN as discussed in chapter 2. The proposed work expects to lessen the energy consumption by deploying multiple sinks [19].

Table 6.1 shows comparison of various parameters of WSN with single sink and multiple sinks. Energy Consumption in single sink WSN is higher because lots of energy is wasted in sending the data from multiple sensing nodes to the single sink. Moreover there is congestion close to the sink. Energy consumption is lower in multiple sinks because traffic will be shared among multiple sinks. The data from the multiple sensing nodes need not travel long distances as the sink is close by. The congestion at one single sink is also avoided. The end to end delay is high in single sink where as quite low in multiple sink[[108]]. There is lower connectivity between nodes in case of single sink. Some of the nodes which are far off from the sensor, do not form the part of the network. But, in the case of multiple sinks, higher connectivity is ensured as the sinks are close to the nodes. The possibility of outliers is also reduced. There is lower data delivery due to congestion in single sink where as in multiple sinks, as the data collecting nodes are high, traffic will be distributed among

Table 6.1.: Comparisons between single sink and multiple sink in WSN

Parameter	Single Sink	Multiple Sink
Energy Consumption	Higher	Lower
End to end delay	Higher	Lower
Connectivity	Lower	Higher
Data Delivery	Lower	Higher
Scalability	Lower	Higher
Data Aggregation	Required	Not Required
Cost	Lower	Higher

different sinks so more data can be successfully delivered to the sinks Sensor nodes have limited transmission range so with single sink, nodes closer to the sink, have to forward lots of packets including the ones from other nodes as well. So the chances to such nodes dying out are increased causing a communication link failure. Whereas, in case of multiple sinks, large scale network can be handled because there is no congestion on a single node closer to the sink. Every node is transmitting data directly to the sink. In single sink the data captured at every CH needs to be aggregated by the CH and then sent to the sink. Lots of energy is consumed in data aggregation. Whereas, in the case of multiple sinks, data need not to be aggregated by the sinks and the energy consumed in data aggregation is saved.

The section 6.2 deals with information about multiple sinks. The section 6.3 gives the proposed algorithm. The sections 6.4 and 6.5 discuss the process of finding the optimal number of sinks and finding the sink locations respectively. The section 6.6 deal with the implementation and results.

6.2. Multiple Sinks

Generally a WSN is based on many-to-one communication concept of having many sensors transferring the data to a common single sink. The proposed work demonstrates the importance of multiple sink in a WSN[17]. The most important advantage of having multiple sinks is to shorten the routing path between a sensing node and the sink. If the area to be sensed is very huge and the sensors are randomly deployed, then it happens that even a very effective routing protocol fails. Even if clustering is used, the nodes waste a lot of energy in sending the data to long distances. If the sink or the next CH is within the transmission range of the sensor then the energy consumed is given by equation 6.1

$$E \propto d^2 \tag{6.1}$$

Where d is the distance from source to destination node.

But if the sink or the next CH is not within the transmission range of the sensor then the energy consumed is given by equation 6.2

$$E \propto d^4 \tag{6.2}$$

So, to save energy, it becomes very important to reduce the distance between the node and the sink or the next CH. The current work solves two problems of having multiple sinks in a WSN namely

1. Optimal number of sinks
2. Position of sinks

Multiple sinks deployment makes the WSN robust so that even if one of the sinks die out, other sinks can take the charge and prevent the WSN from failing.

6.3. Proposed Algorithm

After the initial random deployment of the sensors, the position of the sensors is found out either by the GPS or the triangulation method. The transmission range of the sensors is used to compute the coordinates of the sensors in both the mentioned methods. Once the position of the nodes is found out, the clusters are formed. Once the clusters are formed, the optimum number of sinks is calculated. The steps followed in the proposed algorithm are as follows:

1. Randomly deploy the sensing nodes
2. Determine the coordinates of the nodes using triangulation method or GPS
3. Find the optimal number of clusters keeping in mind that all the nodes should be at single hop from the sink.
4. Cluster the entire region to be sensed.
5. The sinks are positioned at the centroid of the bounded polygon formed by the coordinates of the nodes.
6. Once the sinks are setup, the nodes start sending the data to the sink and the sink aggregates and interprets the data depending upon the application for which the WSN is employed.

Assumptions

Some assumptions are:

1. The nodes after deployment are fixed
2. There are multiple nodes and multiple sinks
3. The sinks are fixed whose position is determined by algorithm
4. The area to be sensed is large enough and the nodes are scarcely deployed to reduce the cost of the WSN
5. All the nodes are homogeneous.
6. No two sinks occupy the same place

Table 6.2.: Notation Table

Description	Notation
The region to be sensed	A
A small part of the region	a'
The center of the region	Ca'
Length of the region	l
Breadth of the region	b
The function to form clusters	$Form\_Cluster()$
A set of sensing nodes	N
A node in the set N	n
All the nodes in the region a'	Na'
Distance of node n from Ca'	dCa'
Transmission range of a node n	TR_n
Total number of clusters	cl

The next two sections discuss the process of finding the optimal number of sinks and then finding their positions respectively.

6.4. Optimal Number of Sinks

While having multiple sinks in a WSN, there is a need of finding the optimal number of sinks. Too many sinks will have a trade off on cost of the network. Too few sinks does not solve the problem of energy saving while sending the sensed data from the node to the sink or the next CH. So the optimal number of sinks needs to be found out. Finding the optimal number of sensors is a NP hard problem [153]. If the Euclidean distance is used as the measure for clustering metric then the centroid of the nodes in a cluster is treated as the location of the multiple sinks. In this case the

Total number of sinks = Total number of clusters

Now finding the number of sinks is now translated to finding out the optimum number of clusters. It is proposed that the number of clusters is decided in such a way that all of the nodes are at single hop from the center of the cluster. The notations used are shown in the Table 6.2

The pseudo code for the formation of the clusters is as in Algorithm 6.1

The function $Form\_Cluster(a')$ is a recursive function which will be called by itself again and again until all the nodes are at single hop from the center of the cluster which is determined a priori.

Algorithm 6.1 Form_Cluster

$Form\_Cluster(a')$
{
    if (all values of $n, n \in Na', TR_n \leq dCa'$)
    $break$;
    $else$
    {
        Divide a' into two halves $(a'/2)$
        Find the set $N_{a'/2}$
        $cl = cl + 1$;
        $Form\_Cluster(a'/2)$
    } // end of else condition
}

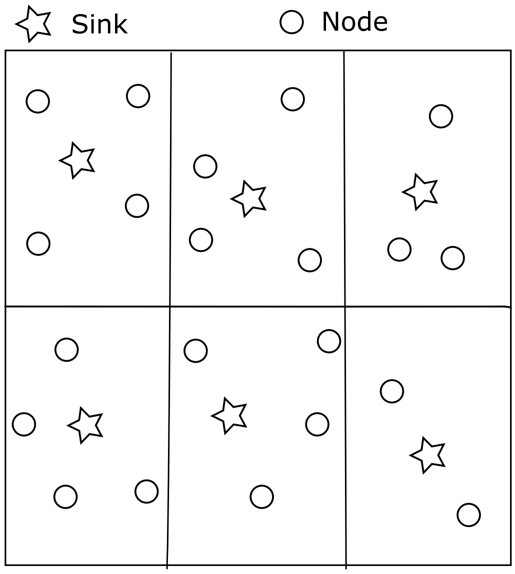

Figure 6.1.: Clusters and sinks in rectangular region

6.5. Finding Sink Locations

It's assumed that the sink is placed at the center of the cluster. So, all the nodes should be at a single hop from the center of the cluster. Single-hop has an edge over multiple-hop in the fact that a lot of energy is wasted in sending the data from one node to another. In multiple-hop, energy of the network is also wasted in finding out the optimum routing path. To minimize the energy consumption in a WSN, single hop clustering is done. In an ideal case for the theoretical explanation we assume that the region to be sensed in a complete rectangle and the rectangle is divided into 6 clusters of equal size. The sinks are placed at the centers of the 6 small rectangles formed within the larger rectangle as shown in the Figure 6.1.

But in real world scenarios, it's not possible that the region is a rectangle. In actuality the region is not uniform. The sensors are deployed randomly as shown in the Figure 6.2.

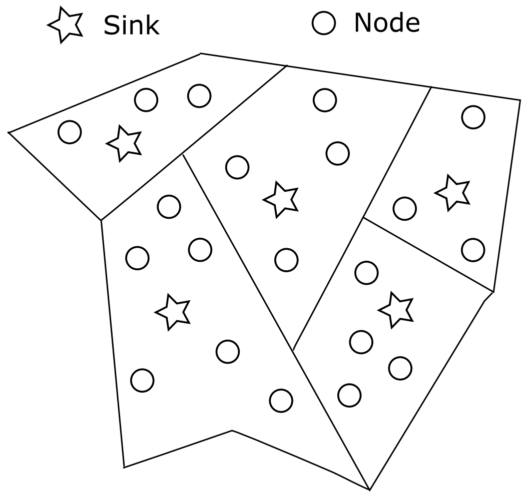

Figure 6.2.: Clusters and sinks in irregular region

The clusters are formed following the pseudo code as explained earlier. The sinks are placed at the centroid of the bounded polygons. The centroid of a bounded polygon is given by

$$C_x = \frac{1}{6A}\sum_{i=0}^{n-1}(x_i + x_{i+1})(x_i y_{i+1} - x_{i+1} y_i) \tag{6.3}$$

$$C_y = \frac{1}{6A}\sum_{i=0}^{n-1}(y_i + y_{i+1})(x_i y_{i+1} - x_{i+1} y_i) \tag{6.4}$$

$$A = \frac{1}{2}\sum_{i=0}^{n-1}(x_i y_{i+1} - x_{i+1} y_i) \tag{6.5}$$

and where A is the polygon's signed area.

In these formulas, the vertices are assumed to be numbered in order of their occurrence along the polygon's perimeter, and the vertex (x_n, y_n) is assumed to be the same as (x_0, y_0). If the points are numbered in clockwise order the area A, computed as above, will have a negative sign; but the centroid coordinates will be correct even in this case.

The sink is placed at the center of the cluster so there is no need to determine the CH. Hence the network energy utilized in CH selection is also preserved.

6.6. Implementation and Simulation Results

The nodes were deployed randomly as shown in the Figure 6.3.

The area is calculated by the Eqn. 6.5. The coordinates of the sink are calculated by the Eqns. 6.3 and 6.4. Then the WSN is simulated on MATLAB and the numbers of rounds are

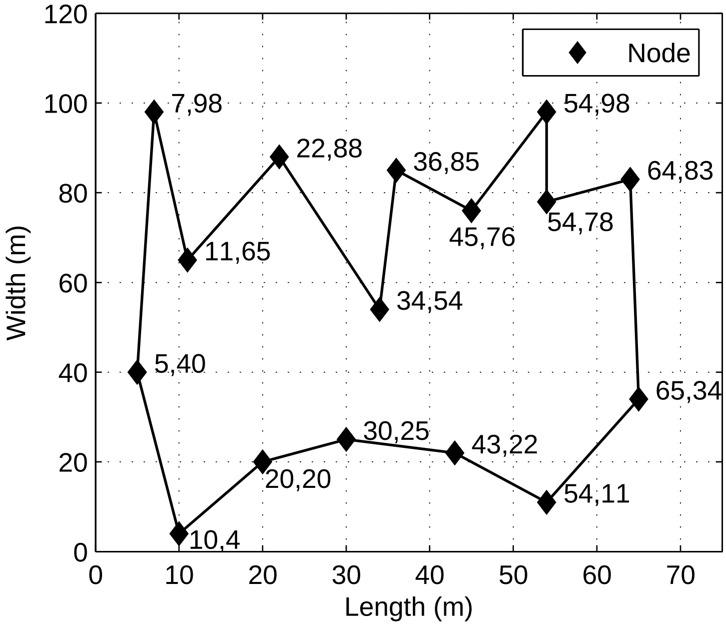

Figure 6.3.: Bounded polygon of randomly deployed nodes

Table 6.3.: Simulation Results

Protocol	Rounds for FND	Rounds for LND
Direct	217	468
MTEs	15	843
Static Clustering	106	240
LEACH	1848	2608
Proposed protocol	3125	5423

found before the first node dies out (FND) and the last node dies out (LND). It is compared with the different existing protocols i.e. Direct, Minimum Transmission Energy (MTE), Static clustering and dynamic clustering (LEACH). The initial energy of each node is assumed to be 1 Joule. The simulation is done with 100 nodes and 10 sinks. The figures for the other protocols are taken from the research paper [3].

The Table 6.3 clearly depicts that the proposed protocol has better energy efficiency as the number of rounds before first node and last node die out is considerably high as compared to other existing protocols [19] . A high improvement is observed for the rounds for FND, and LND as compared to DTE. However, if the results of proposed algorithm are compared with the best existing case, i. e. LEACH case, it is observed that an improvement by 1277 and 2815 round achieved for FND and LND respectively.

6.7. Summary

From the proposed work it is concluded that the network lifetime is considerably increased by using multiple sinks. In single hop clustering the nodes need to send data to sinks located quite far off and thus waste energy. In multiple hops the energy is wasted in finding out the optimal path and then sending the data to the sinks in multiple hops and thus wasting some amount of energy in every hop. Single hop multiple sink clustering reduces the energy consumption in finding out the optimal routing path and then sending the data to the sink in multiple hops.

7. Lifetime Optimization in Multiple Sinks

7.1. Introduction

In a single sink WSN, the nodes need to send the data through multiple hops. In a large WSN, it becomes quite inefficient in terms of power consumption while gathering all information in a single sink [111]. Most extreme energy utilization happens in conveying the information from the nodes to the sink [13, 16]. To minimize the energy utilization while sending the information to the sink, various sinks are utilized. As there are multiple sinks, the separation from the nodes to the sink decreases, in this way there is no need of multiple hops. Multiple sinks decrease the distance the detected information needs to travel and henceforth correspondingly decrease the energy utilization considerably[19]. Another detriment of a solitary sink WSN is that of energy disbalance between the nodes near the sink and the ones which are far away [20]. The system is re-structured by altering the quantity of nodes associated with a sink. In this chapter, a calculation is proposed for system rebuilding in a multiple sink WSN to decrease the energy utilization and expand the system lifetime. This energy adjustment through system re-organization advances the system lifetime. The quantity of not connected notes are additionally entirely less. The execution is done in MATLAB. The execution results support the stated proclamations.

The next section describes the proposed algorithm. The simulation results are discussed in the section 7.3. The chapter is summarized in the section 7.4.

7.2. Proposed Algorithm

The proposed work concentrates on multiple sink single hop routing[21]. The nodes and the sinks are randomly deployed. At the first instance the nodes are connected to a sink depending upon their distance and the transmission energy. A node gets connected to a sink/s if its distance from the sink/s is less than the transmitting range. In this way all the nodes are connected to one or many sinks. There may be some nodes which are not connected to any of the sinks as they are deployed quite far from the network and is not within the transmission range of any sink. In the next phase network restructuring is done. The energy consumed by every sink is calculated and the sink with maximum energy consumption is identified. The unique nodes connected to this sink are traced. A unique node is one which is connected to only that sink. All the other connected nodes apart from the unique nodes are then found out. Such nodes are then connected to other sinks (within the transmission range); keeping in mind that the new energy consumption of that sinks doesn't cross the threshold. In this way the energy consumption of the sink which was earlier consuming maximum energy, is reduced. This process is repeated for all the sinks in the increasing order of their energy consumption. The end result is a network which now consumes less energy overall.

7.2.1. System Model and Assumptions

1. Sinks are randomly deployed and then they are fixed. Since random distribution is used, the complexity in determining the position of the sink is removed.

2. The nodes after random deployment are fixed.

3. The density of nodes deployed is high such that the data reaching a sink in single hop consumes small energy.

4. The network is heterogeneous. The sinks have more power than the sensing nodes. The sinks have additional computational capacity as well.

7.2.2. Pseudo Code

1. The sensor nodes and the sinks are randomly deployed and after the deployment the nodes and sinks are stationary. Combination of sink and sensor nodes will make the network heterogeneous.

2. N is the set of p nodes deployed in the area to be sensed in the given network. $N = \{n_1, n_2, n_3, \ldots\ldots\ldots\ldots n_p\}$

3. S is the set of q sinks deployed in the area to be sensed in the given network. $S = \{S_1, S_2, S_3, \ldots\ldots\ldots\ldots S_q\}$.

4. Calculate the Euclidean distance from each sink to every node. $DS_i$is the set of distances of all the nodes from the i^{th} sink.

$$D_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$

$$i = sink_{ID} = 1, 2, \ldots, q, j = node_{ID} = 1, 2, \ldots, p$$

$$DS_i = \{D_{i1}, D_{i2}, D_{i3}, \ldots\ldots\ldots\ldots D_{ip}\}$$

So it will form a [q, p] order distance matrix (D) which will contain the distances of all the sinks from all the nodes. The distance between the nodes is calculated using a method based on RSSI [154].

5. The threshold energy of the sink is E_0.

6. The transmission range of a node is T_x.

7. The neighboring nodes of every sink are calculated based on the transmitting range. Nb_i is the set of all the neighboring nodes of i^{th} sink (S_i)

$$Nb_i \subseteq N \, where \, \{Nb_i \mid D_{ij} < T_x \in j, j \in N\}$$

8. A new connection matrix (C) is formed based on D. A flag is set for every element where the distance from node to sink is less than the transmission range Tx. Thus C is in the binary form.

9. The energy consumed by i^{th} sink E_i is calculated by

$$E_i = k \sum_{j=1}^{|Nb_i|} (D_{ij})^2$$

Where $\mid Nb_i \mid$ is the total number of neighbor nodes of i^{th} sink (S_i) and D_{ij} is the distance of i^{th} sink from the j^{th} node where $n_j \in Nb_i$, k is the constant for first order radio energy model [3]

10. Calculate the E_{max}= maximum(E_i), Where $i = 1 \, to \, q$ and find the maximum energy consumed by any sink S_i.

11. $if(E_0 > E_{max})$ { No need to optimize the network, Iteration = 0; Set the E_0 below the E_{max}, Repeat step 10}

12. *else*

    a) Calculate the unique nodes connected to a sink. A unique node to a sink is the one which is not connected to any other sink. U_i is the set of unique nodes for $S_i, U_i \in Nb_i$,

    $$U_i = \{n_1, n_2, n_3, \ldots, n_i\} \in Nb_i \& U_i \notin Nb_j$$

    $Where \, j = 1, 2, 3 \ldots, q \, and \, i \neq$ j

Table 7.1.: Simulations Parameters

Parameter	Value	Parameter	Value
N	$50,100$	E_{max}	$Max\,Energy\,consumed\,by$ sink
R_s	$Point\,Sensor$	n_i	$node\,id\,of\,i^{th}\,node$
R_t	$150m$	S	$10, Number\,of\,sinks$
A	$100*100m^2$	E_r	$E_0 - E_{DA} - E_{processing} - E_{sense}$
E_0	$Initial\,1J$	E_{th}	$0.5J\,to\,0.1J$

b) Based on the above step and 7 we can easily calculate the nodes, having connectivity with more than one sink. MC_i is the multiple connecting nodes set having the connection with multiple sinks.

c) MC_i nodes of the i^{th} sink are arranged into the descending order of the distance from the i^{th}sink.

d) Select the nodes having the minimum distance from the i^{th} sink and disconnect the connection of remaining nodes those are far from the i^{th} sink, and update the overall connection matrix C based on the distance matrix D.

13. Repeat the steps 9-12 by re-calculating the sink energy with modified C. The iteration count is also increased.

14. The steps are repeated until E_{max} becomes constant. Now the network is optimized and the routing is started by the nodes. The lifetime of the network is calculated by counting the number of rounds done by the network before the first node die out.

7.3. Simulation and Results

The simulation of the above mentioned pseudo code was performed in MATLAB. The simulation parameters are mentioned in Table 7.1.

The numbers of nodes not connected to any sink as the number of sinks increase in the network are depicted in Table 7.2. In the proposed work we have considered the restructuring energy in terms of processing energy($E_{processing}$). In the restructuring only the node connection changes with one sink to other sink to reduce the load of the sink and decrease the delay.

It's apparent from the Table 7.2, that as we increase the sinks, the number of not connected nodes decreases. But after a certain point of time, the not connected nodes more or less remain same. Since the sinks and the nodes are randomly deployed, and the nodes are connected to the sink with a single hop, there are not connected nodes because of large transmission energy required outside the transmission range $R_t(E \propto d^4)$ [3]. If we would

Table 7.2.: Before Network Re-structuring

Sinks	*Not Connected*	E_{max} (J)	E_{min} (J)	E_{avg} (J)	*Max nodes connected to single sink*
5	30	0.312	0.173	0.240	28
6	28	0.331	0.139	0.23	30
7	22	0.359	0.152	0.244	31
8	16	0.371	0.114	0.239	31
9	14	0.334	0.121	0.227	30
10	11	0.362	0.103	0.233	32
11	10	0.360	0.1	0.231	31
12	9	0.386	0.110	0.236	32
13	10	0.363	0.101	0.23	30
14	9	0.368	0.104	0.226	31
15	8	0.391	0.093	0.226	32

Table 7.3.: After Network Re-structuring

Sinks	E_{max} (J)	E_{min} (J)	E_{avg} (J)	*Max nodes connected to single sink*
5	0.240	0.173	0.195	15
6	0.230	0.139	0.193	15
7	0.244	0.152	0.192	10
8	0.247	0.114	0.211	12
9	0.288	0.121	0.210	11
10	0.274	0.103	0.20	12
11	0.245	0.121	0.197	10
12	0.245	0.110	0.166	8
13	0.260	0.101	0.194	7
14	0.240	0.104	0.198	6
15	0.242	0.093	0.217	7

have deployed the sinks manually, the number of not connected nodes would have reduced considerably. The optimal number of sinks also depends upon the network area to be covered by the sinks.

The Figure 7.1 depicts the plot of total percentage of sinks vs. percentage of not connected nodes. With the help of curve fitting tool of 4^{th} order polynomial obtained is

$$y = -0.0143x^4 + 0.5585x^3 - 7.4997x^2 + 37.6059x - 30.9394 \tag{7.1}$$

Using the above polynomial, suitable value is 0.12 sink/nodes. This shows that the optimal number of sinks is 12 for 100 nodes in the considered scenario. Now, the network restructuring algorithm mentioned in the pseudo code section is applied to the above data as depicted in Table 7.3.

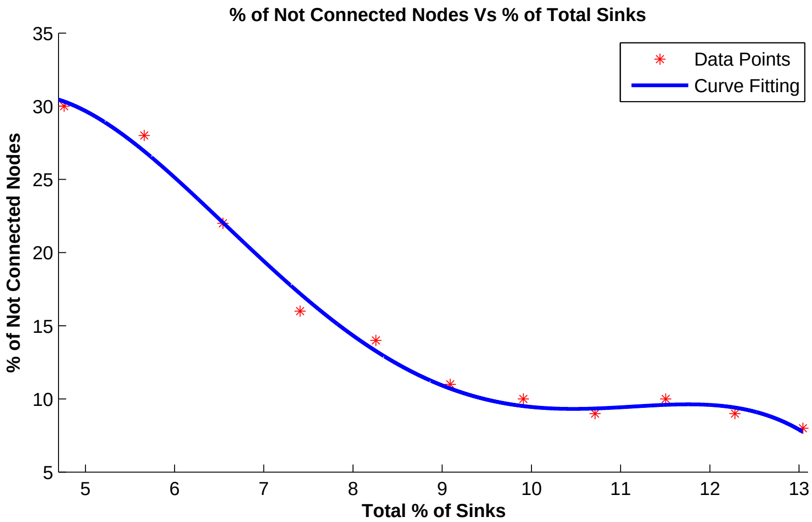

Figure 7.1.: Not connected nodes

Table 7.4.: Before Network Re-structuring Sinks = 10

Network Restructuring	*Sinks*	E_{max} (J)	E_{min} (J)	E_{avg} (J)	*Max nodes connected to single sink*
Before	10	0.362	0.103	0.233	32
After	10	0.274	0.103	0.20	12

It can be observed that after restructuring, the network consumes less energy. The number of maximum nodes connected to single sink reduces which results in reduction of the total energy consumption of sink. Thus, it results in increased network lifetime.E_{min} remains the same before and after the network restructuring because the number of nodes connected to particular sink is minimum and constant. Where as E_{max} is reduced due to network restructuring. If a node is connected to multiple sinks then after restructuring, it gets connected to single sink to save the other sinks energy consumption. So, E_{avg} is automatically decreased.

In Table 7.4, we consider the case of network with 10 Sinks. After applying the algorithm, i.e. after energy balancing through network restructuring

The E_{max} has changed from 0.362 to 0.274. The maximum number of nodes connected to a single sink has been changed to 32 to 12. The E_{avg} is also changed from 0.233 to 0.20.

To appreciate the algorithm, a Table7.5 is added which gives the change in values of E_{max} before and after network restructuring and also mentions the percentage decrease in E_{max}.

The Table 7.5 clearly shows that there is considerable decrease in the E_{max} before and after network restructuring.

A similarTable 7.6 is shown to depict the percentage decrease in the maximum number of nodes connected to a single sink after network restructuring.

The Table 7.6clearly shows that there is considerable decrease in the maximum number of

Table 7.5.: Percentage decrease in E_{max}

Number of sinks	E_{max}		
	Before restructuring	After restructuring	% decrease
5	0.312	0.240	23.1%
6	0.331	0.230	30.5%
7	0.359	0.244	32.0%
8	0.371	0.247	33.4%
9	0.334	0.288	13.8%
10	0.362	0.274	24.3%
11	0.360	0.245	31.9%
12	0.386	0.245	36.5%
13	0.363	0.260	28.4%
14	0.368	0.240	34.8%
15	0.391	0.242	38.1%

Table 7.6.: Percentage decrease in the maximum number of nodes

Number of sinks	Maximum number of nodes connected to a single sink		% decrease
	Before the restructuring	After the restructuring	
5	28	15	46.4%
6	30	15	50.0%
7	31	10	67.7%
8	31	12	61.2%
9	30	11	63.3%
10	32	12	62.5%
11	31	10	67.7%
12	32	8	75.0%
13	30	7	76.7%
14	31	6	80.6%
15	32	7	78.1%

Table 7.7.: Maximum nodes connected to single sink

No. of	Maximum number of nodes connected to a single sink	
Nodes	Before the restructuring	After the restructuring
80	30	11
90	30	11
100	32	12
110	36	14
120	35	13
130	40	14

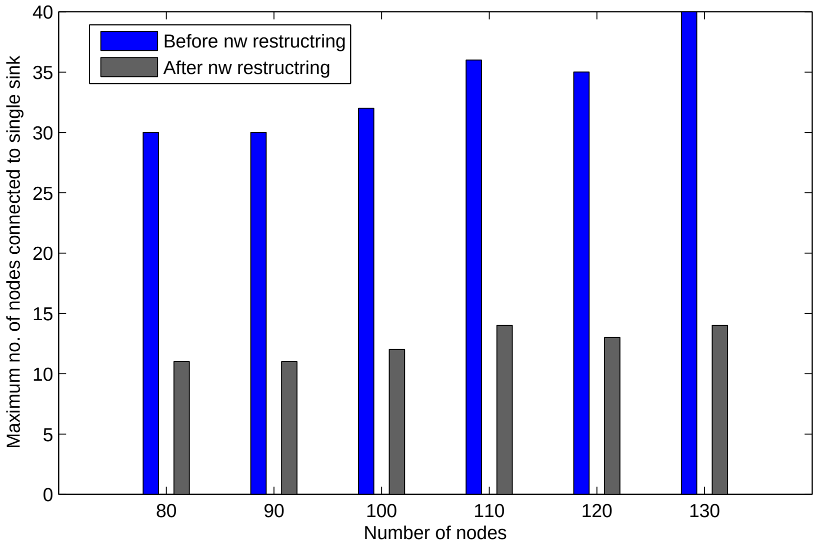

Figure 7.2.: Maximum nodes connected to single sink

nodes connected to a single sink after network restructuring. E_{avg} is also decreased, a table for which is not shown separately.

Now, we consider one more scenario of changing the number of sensing nodes. The observations are depicted in Table 7.7, Figure 7.2.

Considering the Table 7.7, when the number of nodes are 80, the maximum number of nodes connected to a single sink before and after network restructuring has been reduced to 11 from 30. The percentage drop in the maximum number of nodes connected to a single sink is depicted in the Table 7.8

Thus it shows that the proposed algorithm is quite effective for multiple sink, single hop WSN. After the network restructuring, the maximum number of nodes connected to a sink reduces considerably. As a result of this energy balancing, the network lifetime is increased and the energy consumption is reduced.

The optimal number of sinks is also calculated using fourth order polynomial. The number of iterations used for simulation is 100. In 100 iterations the value of average energy (mean)

Table 7.8.: Percentage drop in maximum number of nodes

No of nodes	Percentage drop in max no. of nodes connected to a single sink
80	63.3%
90	63.3%
100	62.5%
110	61.1%
120	62.9%
130	65%

is $0.2329J$ and the standard deviation of mean is $0.0061J$. Since the standard deviation is very small so this iteration count is statistically valid.

7.4. Summary

The proposed work performs network lifetime optimization through energy balancing in a multiple sink single hop WSN. The network restructuring balances the energy among the sinks, thereby increasing the network lifetime. In the proposed algorithm we have considered the network lifetime with respect to the maximum and average energy consumption. More the energy consumed and less the network lifetime. The proposed restructuring algorithm reduces the maximum nodes connected to a sink. As a result of which, the total energy consumed by the sink connected to the maximum nodes also decreases further resulting in increasing the total network lifetime. The implementation results shown in MATLAB demonstrate that network restructuring is beneficial in reducing the maximum and average energy of the sinks. The optimal number of sinks are also calculated through a fourth order polynomial.

8. Conclusion

A wireless sensor network consists of a large number of sensing nodes randomly deployed in a region. The objective of the research was to design a WSN so as to achieve the following:-

1. Maximum energy utilization

2. Maximum Connectivity and coverage of the network

3. Optimized Network Lifetime of the network.

In this chapter, the contributions of the thesis have been summarized and some suggestions for future work have also been provided.

8.1. Summary of Contributions

The current research work proposes a WSN for sensing multiple attributes in an environment and then producing an alarm if any unexpected condition is predicted. The protocol used is hybrid and bidirectional. The use of data aggregation has minimized the data transmissions and also made it fast. The parallel processing in the application processor has fastened up the data analysis as depicted in the chapter 3 of the current thesis work.

We have also proposed a merging algorithm which minimizes the global energy usage by making sensing nodes to work in sleep and active mode. Merging outperforms the conventional network's energy usage by uniformly distributing the load of sensing in the nodes falling in the same sensing range. Clearly our simulations as depicted in chapter 3, show that

- Network using merging is more energy efficient than the conventional network.

- Lifetime of network using merging is increased.

The application of this type of network become limited in sparse deployment as the sensing range of the nodes is less. There has to be large number of nodes for feasibility of this network. Based on our MATLAB simulations as mentioned in the section 3.4.4, our proposed model outperforms the conventional models. Providing WSNs with such efficient models will open up a whole new horizon for them.

Clustering in WSN not only reduces energy consumption but also achieves better network performance. Hierarchical agglomerative clustering uses bottom up approach and hence the complete network information is not required. The dendrograms for the various HAC protocols for qualitative and quantitative data are shown in the chapter 4. If the position of the nodes is not known due to the issues like cost, time, energy conservation then the Sorensen method is used to find out the similarity matrix and then perform the clustering of the nodes. The dendrograms for single and complete link protocol clearly show that single link has a chain effect, making the clustering inefficient. The dendrograms of quantitative single link, complete link, and average link are similar to the dendrograms of qualitative single link, complete link and average link respectively. So, it shows, as in section 4.4, that there is no need to find out the location of the nodes in a randomly deployed WSN. Finding out the position of the nodes not only make the clustering complex but also consumes a lot of energy.

The most energy consuming process in clustering is the determination of CHs. The proposed algorithm for CH selection in a WSN using unique node concept has many advantages as shown in the chapter 5. The latency in transmitting the data in a single hop is much more than in the proposed multi hop WSN. Each node in the network transmits the data to CH/BS closest to it. The CH in turn transmits the data to the next CH, if required, to reach the BS. If the CH is selected on the basis of the concept of maximum number of nodes connected, then it may happen that one or more unique nodes are not connected to any of the selected CHs. Thus this algorithm deals with the CH selection based on the unique node concept. In the proposed algorithm there is no possibility of having any outlier, as all the unique nodes are connected to some or the other CH. Adaptability is also well taken care of. After clustering, if in the process of transmission of data, one of the CHs die out, then the CH at the previous hop comes to know about it since the data from the died CH didn't reach it and the clustering algorithm is repeated after $(n-1)^{th}$ hop for the entire network. If the residual energy of the CH becomes less than the threshold energy then the CH section process is re-instantiated after $(n-1)^{th}$ hop for that particular path. The results as depicted in section 5.4, clearly depict that the proposed clustering algorithm increases the network lifetime. In WSNs, the communication cost is often several orders of magnitude larger than the computation cost, thus the CHs perform data aggregation to reduce the amount of data to be transmitted.

The motivation of the current research work is to reduce the energy consumption in a WSN and consequently increase the network lifetime. From the proposed work in the chapter 6, it is concluded that the network lifetime is considerably increased by using multiple sinks. In single hop clustering the nodes need to send data to sinks located quite far off and thus waste energy. In multiple hops the energy is wasted in finding out the optimal path and then sending the data to the sinks in multiple hops and thus wasting some amount of energy in every hop. Single hop multiple sink clustering reduces the energy consumption in finding out the optimal routing path and then sending the data to the sink in multiple hops.

The proposed work in chapter 7,performs network lifetime optimization through energy balancing in a multiple sink single hop WSN. The network restructuring balances the energy among the sinks, thereby increasing the network lifetime. The proposed restructuring algorithm reduces the maximum nodes connected to a sink. As a result of which the total energy consumed by the sink connected to the maximum nodes also decreases further resulting in increasing the total network lifetime. The implementation results shown in MATLAB as in section 7.3 demonstrate that network lifetime is increased by network restructuring.

8.2. Future Work

While developing the Wireless Enviormntal Monitoring System(WEMS), the future work corresponds to dealing with a hybrid bidirectional protocol for tree based WSN. Then have a comparison between the two approaches based on energy consumption and time delay. In the future the authors wish to consider the scenario of a mobile sinks WSN. In this case there would be single hop routing with the sink moving randomly. This involves complexity of finding out the position of the sinks. The authors also wish to have sinks deployed at fixed positions in the initial phase instead of random deployment. And then finally compare the results obtained in different scenarios with multiple sink moving/stationary with and without initial sink positioning.

In the future the authors plan to propose an algorithm for controlling the traffic at one sink. In a multiple sink WSN, it may happen that the traffic at one sink is very high as compared to the traffic at other sinks. There is a need to monitor this traffic and distribute it to other sinks. Also the authors wish to study the pros and cons of having mobile multiple sinks.

References

[1] Ian F Akyildiz, Weilian Su, Yogesh Sankarasubramaniam, and Erdal Cayirci. A survey on sensor networks. *Communications magazine, IEEE*, 40(8):102–114, 2002.

[2] Azad Ali, Abdelmajid Khelil, Faisal Karim Shaikh, and Neeraj Suri. Efficient predictive monitoring of wireless sensor networks. *International Journal of Autonomous and Adaptive Communications Systems*, 5(3):233–254, 2012.

[3] Wendi Rabiner Heinzelman, Anantha Chandrakasan, and Hari Balakrishnan. Energy-efficient communication protocol for wireless microsensor networks. In *System sciences, 2000. Proceedings of the 33rd annual Hawaii international conference on*, pages 10–pp. IEEE, 2000.

[4] Tapan Jain. Wireless environmental monitoring system (wems) using data aggregation in a bidirectional hybrid protocol. In *Information Systems, Technology and Management*, pages 414–420. Springer, 2012.

[5] Errol L Lloyd and Guoliang Xue. Relay node placement in wireless sensor networks. *Computers, IEEE Transactions on*, 56(1):134–138, 2007.

[6] Jamal N Al-Karaki and Ahmed E Kamal. Routing techniques in wireless sensor networks: a survey. *Wireless communications, IEEE*, 11(6):6–28, 2004.

[7] A. Arman, V. Kher, T. Jain, and D.S. Saini. Implementation of mpls routing in traditional leach. In *Signal Processing and Integrated Networks (SPIN), 2015 2nd International Conference on*, pages 1026–1030, Feb 2015.

[8] Levente Buttyán and Tamás Holczer. Private cluster head election in wireless sensor networks. In *Mobile Adhoc and Sensor Systems, 2009. MASS'09. IEEE 6th International Conference on*, pages 1048–1053. IEEE, 2009.

[9] Apeksha Chauhan, Raghove Bhargove, Tanvi Sharma, and Tapan Kumar Jain. A framework for energy efficient routing protocol for homogeneous wireless sensor networks using sensing range. In *Signal Processing and Integrated Networks (SPIN), 2014 International Conference on*, pages 438–443. IEEE, 2014.

[10] Hoda Taheri, Peyman Neamatollahi, Mohammad Hossein Yaghmaee, and Mahmoud Naghibzadeh. A local cluster head election algorithm in wireless sensor networks. In *Computer Science and Software Engineering (CSSE), 2011 CSI International Symposium on*, pages 38–43. IEEE, 2011.

[11] Jin-Su Kim, Seong-Yong Choi, Seung-Jin Han, Jun-Hyeog Choi, Jung-Hyun Lee, and Kee-Wook Rim. Alternative cluster head selection protocol for energy efficiency in wireless sensor networks. In *Future Dependable Distributed Systems, 2009 Software Technologies for*, pages 159–163. IEEE, 2009.

[12] Sang H Kang and Thinh Nguyen. Distance based thresholds for cluster head selection in wireless sensor networks. *Communications Letters, IEEE*, 16(9):1396–1399, 2012.

[13] Tapan Kumar Jain, Davinder Singh Saini, and Sunil Vidya Bhooshan. Cluster head selection in a homogeneous wireless sensor network ensuring full connectivity with minimum isolated nodes. *Journal of Sensors*, 2014, 2014.

[14] Tapan Kumar Jain, Davinder S Saini, and Sunil V Bhooshan. Performance analysis of hierarchical agglomerative clustering in a wireless sensor network using quantitative data. In *Information Systems and Computer Networks (ISCON), 2014 International Conference on*, pages 99–104. IEEE, 2014.

[15] Jing Teng, Hichem Snoussi, Cédric Richard, and Rong Zhou. Distributed variational filtering for simultaneous sensor localization and target tracking in wireless sensor networks. *Vehicular Technology, IEEE Transactions on*, 61(5):2305–2318, 2012.

[16] Deepak R Dandekar and PR Deshmukh. Energy balancing multiple sink optimal deployment in multi-hop wireless sensor networks. In *Advance Computing Conference (IACC), 2013 IEEE 3rd International*, pages 408–412. IEEE, 2013.

[17] Pietro Ciciriello, Luca Mottola, and Gian Pietro Picco. Efficient routing from multiple sources to multiple sinks in wireless sensor networks. In *Wireless Sensor Networks*, pages 34–50. Springer, 2007.

[18] Theodore S Rappaport et al. *Wireless communications: principles and practice*, volume 2. prentice hall PTR New Jersey, 1996.

[19] Tapan Kumar Jain, DS Saini, and SV Bhooshan. Increasing lifetime of a wireless sensor network using multiple sinks. In *Information Technology: New Generations (ITNG), 2014 11th International Conference on*, pages 616–619. IEEE, 2014.

[20] Zhao Cheng, Mark Perillo, and Wendi B Heinzelman. General network lifetime and cost models for evaluating sensor network deployment strategies. *Mobile Computing, IEEE Transactions on*, 7(4):484–497, 2008.

[21] Tapan Kumar Jain, Davinder Singh Saini, and Sunil Vidya Bhooshan. Lifetime optimization of a multiple sink wireless sensor network through energy balancing. *Journal of Sensors*, 2015, 2015.

[22] Hannes Frey, Stefan Rührup, and Ivan Stojmenović. Routing in wireless sensor networks. In *Guide to Wireless Sensor Networks*, pages 81–111. Springer, 2009.

[23] Liliana M.C. Arboleda and Nidal Nasser. Comparison of clustering algorithms and protocols for wireless sensor networks. In *Electrical and Computer Engineering, 2006. CCECE '06. Canadian Conference on*, pages 1787–1792, May 2006.

[24] Asis Kumar Tripathy and Suchismita Chinara. Comparison of residual energy-based clustering algorithms for wireless sensor network. *ISRN Sensor Networks*, 2012, 2012.

[25] Rajashree V Biradar, VC Patil, SR Sawant, and RR Mudholkar. Classification and comparison of routing protocols in wireless sensor networks. *Ubicc journal*, 4:704–711, 2009.

[26] Laiali Almazaydeh, Eman Abdelfattah, Manal Al-Bzoor, and Amer Al-Rahayfeh. Performance evaluation of routing protocols in wireless sensor networks. *International Journal of Computer Science and Information Technology*, 2(2):64–73, 2010.

[27] Samira Yessad, Louiza Bouallouche-Medjkoune, and Djamil Aıssani. Comparison of routing protocols in wireless sensor networks.

[28] Shio Kumar Singh, MP Singh, DK Singh, et al. Routing protocols in wireless sensor networks–a survey. *International Journal of Computer Science & Engineering Survey (IJCSES) Vol*, 1:63–83, 2010.

[29] Ya Xu, John Heidemann, and Deborah Estrin. Geography-informed energy conservation for ad hoc routing. In *Proceedings of the 7th annual international conference on Mobile computing and networking*, pages 70–84. ACM, 2001.

[30] Yan Yu, Ramesh Govindan, and Deborah Estrin. Geographical and energy aware routing: A recursive data dissemination protocol for wireless sensor networks. Technical report, Technical report ucla/csd-tr-01-0023, UCLA Computer Science Department, 2001.

[31] Benjie Chen, Kyle Jamieson, Hari Balakrishnan, and Robert Morris. Span: An energy-efficient coordination algorithm for topology maintenance in ad hoc wireless networks. *Wireless networks*, 8(5):481–494, 2002.

[32] Badri Nath and Dragoş Niculescu. Routing on a curve. *ACM SIGCOMM Computer Communication Review*, 33(1):155–160, 2003.

[33] Guoliang Xing, Chenyang Lu, Robert Pless, and Qingfeng Huang. On greedy geographic routing algorithms in sensing-covered networks. In *Proceedings of the 5th ACM international symposium on Mobile ad hoc networking and computing*, pages 31–42. ACM, 2004.

[34] Michele Zorzi and Ramesh R Rao. Geographic random forwarding (geraf) for ad hoc and sensor networks: multihop performance. *Mobile Computing, IEEE Transactions on*, 2(4):337–348, 2003.

[35] Volkan Rodoplu and Teresa H Meng. Minimum energy mobile wireless networks. *Selected Areas in Communications, IEEE Journal on*, 17(8):1333–1344, 1999.

[36] Li Li and Joseph Y Halpern. Minimum-energy mobile wireless networks revisited. In *Communications, 2001. ICC 2001. IEEE International Conference on*, volume 1, pages 278–283. IEEE, 2001.

[37] Nirupama Bulusu, John Heidemann, and Deborah Estrin. Gps-less low-cost outdoor localization for very small devices. *Personal Communications, IEEE*, 7(5):28–34, 2000.

[38] Wendi Rabiner Heinzelman, Joanna Kulik, and Hari Balakrishnan. Adaptive protocols for information dissemination in wireless sensor networks. In *Proceedings of the 5th annual ACM/IEEE international conference on Mobile computing and networking*, pages 174–185. ACM, 1999.

[39] Joanna Kulik, Wendi Heinzelman, and Hari Balakrishnan. Negotiation-based protocols for disseminating information in wireless sensor networks. *Wireless networks*, 8(2/3):169–185, 2002.

[40] Chalermek Intanagonwiwat, Ramesh Govindan, and Deborah Estrin. Directed diffusion: a scalable and robust communication paradigm for sensor networks. In *Proceedings of the 6th annual international conference on Mobile computing and networking*, pages 56–67. ACM, 2000.

[41] David Braginsky and Deborah Estrin. Rumor routing algorthim for sensor networks. In *Proceedings of the 1st ACM international workshop on Wireless sensor networks and applications*, pages 22–31. ACM, 2002.

[42] Yong Yao and Johannes Gehrke. The cougar approach to in-network query processing in sensor networks. *ACM Sigmod Record*, 31(3):9–18, 2002.

[43] Narayanan Sadagopan, Bhaskar Krishnamachari, and Ahmed Helmy. The acquire mechanism for efficient querying in sensor networks. In *Sensor Network Protocols and Applications, 2003. Proceedings of the First IEEE. 2003 IEEE International Workshop on*, pages 149–155. IEEE, 2003.

[44] Azzedine Boukerche, Xiuzhen Cheng, and Joseph Linus. Energy-aware data-centric routing in microsensor networks. In *Proceedings of the 6th ACM international workshop on Modeling analysis and simulation of wireless and mobile systems*, pages 42–49. ACM, 2003.

[45] Basilis Mamalis, Damianos Gavalas, Charalampos Konstantopoulos, and Grammati Pantziou. Clustering in wireless sensor networks. *RFID and Sensor Networks: Architectures, Protocols, Security and Integrations, Y. Zhang, LT Yang, J. Chen, eds*, pages 324–353, 2009.

[46] Wang Xiao-yun, Yang Li-zhen, and Chen Ke-fei. Sleach: Secure low-energy adaptive clustering hierarchy protocol for wireless sensor networks. *Wuhan University Journal of Natural Sciences*, 10(1):127–131, 2005.

[47] Muhamnmad Omer Farooq, Abdul Basit Dogar, and Ghalib Asadullah Shah. Mr-leach: multi-hop routing with low energy adaptive clustering hierarchy. In *Sensor Technologies and Applications (SENSORCOMM), 2010 Fourth International Conference on*, pages 262–268. IEEE, 2010.

[48] G Santhosh Kumar, Vinu Paul, and K Poulose Jacob. Mobility metric based leach-mobile protocol. In *Advanced Computing and Communications, 2008. ADCOM 2008. 16th International Conference on*, pages 248–253. IEEE, 2008.

[49] Md Solaiman Ali, Tanay Dey, and Rahul Biswas. Aleach: Advanced leach routing protocol for wireless microsensor networks. In *Electrical and Computer Engineering, 2008. ICECE 2008. International Conference on*, pages 909–914. IEEE, 2008.

[50] Mu Tong and Minghao Tang. Leach-b: An improved leach protocol for wireless sensor network. In *Wireless Communications Networking and Mobile Computing (WiCOM), 2010 6th International Conference on*, pages 1–4. IEEE, 2010.

[51] Wendi B Heinzelman, Anantha P Chandrakasan, and Hari Balakrishnan. An application-specific protocol architecture for wireless microsensor networks. *Wireless Communications, IEEE Transactions on*, 1(4):660–670, 2002.

[52] Fan Xiangning and Song Yulin. Improvement on leach protocol of wireless sensor network. In *Sensor Technologies and Applications, 2007. SensorComm 2007. International Conference on*, pages 260–264. IEEE, 2007.

[53] Wendi Beth Heinzelman. *Application-specific protocol architectures for wireless networks*. PhD thesis, Massachusetts Institute of Technology, 2000.

[54] Jing Chen and Hong Shen. Meleach-l: More energy-efficient leach for large-scale wsns. In *Wireless Communications, Networking and Mobile Computing, 2008. WiCOM'08. 4th International Conference on*, pages 1–4. IEEE, 2008.

[55] Stephanie Lindsey and Cauligi S Raghavendra. Pegasis: Power-efficient gathering in sensor information systems. In *Aerospace conference proceedings, 2002. IEEE*, volume 3, pages 3–1125. IEEE, 2002.

[56] Ossama Younis and Sonia Fahmy. Distributed clustering in ad-hoc sensor networks: A hybrid, energy-efficient approach. In *INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*, volume 1. IEEE, 2004.

[57] AMADP Agrawal. Teen: a protocol for enhanced efficiency in wireless sensor networks. *IEEE, San Francisco*, 2001.

[58] Arati Manjeshwar and Dharma P Agrawal. Apteen: A hybrid protocol for efficient routing and comprehensive information retrieval in wireless sensor networks. In *Parallel and Distributed Processing Symposium, International*, volume 2, pages 0195b–0195b. IEEE Computer Society, 2002.

[59] Jun Luo and J-P Hubaux. Joint mobility and routing for lifetime elongation in wireless sensor networks. In *INFOCOM 2005. 24th annual joint conference of the IEEE computer and communications societies. Proceedings IEEE*, volume 3, pages 1735–1746. IEEE, 2005.

[60] Rahul C Shah, Sumit Roy, Sushant Jain, and Waylon Brunette. Data mules: Modeling and analysis of a three-tier architecture for sparse sensor networks. *Ad Hoc Networks*, 1(2):215–233, 2003.

[61] Wensheng Zhang, Guohong Cao, and Tom La Porta. Dynamic proxy tree-based data dissemination schemes for wireless sensor networks. *Wireless Networks*, 13(5):583–595, 2007.

[62] Stephanie Lindsey, Cauligi Raghavendra, and Krishna Sivalingam. Data gathering in sensor networks using the energy* delay metric. In *Parallel and Distributed Processing Symposium, International*, volume 3, pages 30188b–30188b. IEEE Computer Society, 2001.

[63] Stephanie Lindsey, Cauligi Raghavendra, and Krishna M. Sivalingam. Data gathering algorithms in sensor networks using energy metrics. *Parallel and Distributed Systems, IEEE Transactions on*, 13(9):924–935, 2002.

[64] Maurice Chu, Horst Haussecker, and Feng Zhao. Scalable information-driven sensor querying and routing for ad hoc heterogeneous sensor networks. *International Journal of High Performance Computing Applications*, 16(3):293–313, 2002.

[65] Xiaojiang Du and Fengjing Lin. Improving routing in sensor networks with heterogeneous sensor nodes. In *Vehicular Technology Conference, 2005. VTC 2005-Spring. 2005 IEEE 61st*, volume 4, pages 2528–2532. IEEE, 2005.

[66] Ian F Akyildiz, Weilian Su, Yogesh Sankarasubramaniam, and Erdal Cayirci. Wireless sensor networks: a survey. *Computer networks*, 38(4):393–422, 2002.

[67] Tian He, John A Stankovic, Chenyang Lu, and Tarek Abdelzaher. Speed: A stateless protocol for real-time communication in sensor networks. In *Distributed Computing Systems, 2003. Proceedings. 23rd International Conference on*, pages 46–55. IEEE, 2003.

[68] David B Johnson and David A Maltz. Dynamic source routing in ad hoc wireless networks. In *Mobile computing*, pages 153–181. Springer, 1996.

[69] Charles E Perkins and Elizabeth M Royer. Ad-hoc on-demand distance vector routing. In *Mobile Computing Systems and Applications, 1999. Proceedings. WMCSA'99. Second IEEE Workshop on*, pages 90–100. IEEE, 1999.

[70] Kemal Akkaya and Mohamed Younis. An energy-aware qos routing protocol for wireless sensor networks. In *Distributed Computing Systems Workshops, 2003. Proceedings. 23rd International Conference on*, pages 710–715. IEEE, 2003.

[71] H.K. Sethi, S. Dash, M.R. Lenka, and A.R. Swain. Incremental model for complete area coverage in wireless sensor networks. In *Computational Intelligence and Networks (CINE), 2015 International Conference on*, pages 92–97, Jan 2015.

[72] Honghai Zhang and Jennifer C Hou. Maintaining sensing coverage and connectivity in large sensor networks. *Ad Hoc & Sensor Wireless Networks*, 1(1-2):89–124, 2005.

[73] Xiaorui Wang, Guoliang Xing, Yuanfang Zhang, Chenyang Lu, Robert Pless, and Christopher Gill. Integrated coverage and connectivity configuration in wireless sensor networks. In *Proceedings of the 1st international conference on Embedded networked sensor systems*, pages 28–39. ACM, 2003.

[74] D. Bajaj and Manju. Maximum coverage heuristics (mch) for target coverage problem in wireless sensor network. In *Advance Computing Conference (IACC), 2014 IEEE International*, pages 300–305, Feb 2014.

[75] Hongwu Zhang, Hongyuan Wang, and Hongcai Feng. A distributed optimum algorithm for target coverage in wireless sensor networks. In *Information Processing, 2009. APCIP 2009. Asia-Pacific Conference on*, volume 2, pages 144–147. IEEE, 2009.

[76] Xiaole Bai, Ziqiu Yun, Dong Xuan, Ten H Lai, and Weijia Jia. Optimal patterns for four-connectivity and full coverage in wireless sensor networks. *Mobile Computing, IEEE Transactions on*, 9(3):435–448, 2010.

[77] Jennifer C Hou, David KY Yau, Chris YT Ma, Yong Yang, Honghai Zhang, I-Hong Hou, Nageswara SV Rao, and Mallikarjun Shankar. Coverage in wireless sensor networks. In *Guide to Wireless Sensor Networks*, pages 47–79. Springer, 2009.

[78] Ruiying Li, Xiaoxi Liu, Wei Xie, and Ning Huang. Deployment-based lifetime optimization model for homogeneous wireless sensor network under retransmission. *Sensors*, 14(12):23697–23723, 2014.

[79] M. Gupta, C.R. Krishna, and D. Prasad. Seeds: Scalable energy efficient deployment scheme for homogeneous wireless sensor network. In *Issues and Challenges in Intelligent Computing Techniques (ICICT), 2014 International Conference on*, pages 416–423, Feb 2014.

[80] Guan Zhi-yan and Wang Jian-zhen. Research on coverage and connectivity for heterogeneous wireless sensor network. In *Computer Science Education (ICCSE), 2012 7th International Conference on*, pages 1239–1242, July 2012.

[81] Mihaela Cardei and Jie Wu. Energy-efficient coverage problems in wireless ad-hoc sensor networks. *Computer communications*, 29(4):413–420, 2006.

[82] Chi-Fu Huang, Yu-Chee Tseng, and Li-Chu Lo. The coverage problem in three-dimensional wireless sensor networks. *Journal of Interconnection Networks*, 8(03):209–227, 2007.

[83] Joseph o'Rourke. *Computational geometry in C*. Cambridge university press, 1998.

[84] Andrew Howard, Maja J Matarić, and Gaurav S Sukhatme. An incremental self-deployment algorithm for mobile sensor networks. *Autonomous Robots*, 13(2):113–126, 2002.

[85] Anthony Man-Cho So and Yinyu Ye. On solving coverage problems in a wireless sensor network using voronoi diagrams. In *Internet and Network Economics*, pages 584–593. Springer, 2005.

[86] Raymond Mulligan and Habib M Ammari. Coverage in wireless sensor networks: A survey. *Network Protocols and Algorithms*, 2(2):27–53, 2010.

[87] Sasha Slijepcevic and Miodrag Potkonjak. Power efficient organization of wireless sensor networks. In *Communications, 2001. ICC 2001. IEEE International Conference on*, volume 2, pages 472–476. IEEE, 2001.

[88] Qun Zhao and Mohan Gurusamy. Lifetime maximization for connected target coverage in wireless sensor networks. *IEEE/ACM Transactions on Networking (TON)*, 16(6):1378–1391, 2008.

[89] Albert Krohn, Michael Beigl, Christian Decker, Till Riedel, Tobias Zimmer, and David Garces. Increasing connectivity in wireless sensor network using cooperative transmission. In *3rd International Conference on Networked Sensing Systems (INSS)*. Chicago, USA, 2006.

[90] Xiaojiang Du and Hsiao-Hwa Chen. Security in wireless sensor networks. *Wireless Communications, IEEE*, 15(4):60–66, 2008.

[91] Adrian Perrig, John Stankovic, and David Wagner. Security in wireless sensor networks. *Communications of the ACM*, 47(6):53–57, 2004.

[92] Chili-Wen Chen, Kuo-Feng Ssu, and Hewijin Christine Jiau. Fault-tolerant topology control with adjustable transmission ranges in wireless sensor networks. In *Dependable Computing, 2007. PRDC 2007. 13th Pacific Rim International Symposium on*, pages 131–138. IEEE, 2007.

[93] S.S. Baidya and A. Baidya. Energy conservation in a wireless sensor network by an efficient routing mechanism. In *Communication, Information Computing Technology (ICCICT), 2015 International Conference on*, pages 1–6, Jan 2015.

[94] Felipe Denis Mendonca Oliveira, Rodrigo Soares Semente, Jefferson Doolan Fernandes, Talison Augusto Correia Melo, and Andres Ortiz Salazar. Seree: An energy-efficient wireless sensor network embedded system to be applied on plunger lift oil elevation method. *Latin America Transactions, IEEE (Revista IEEE America Latina)*, 13(4):1187–1197, April 2015.

[95] John Thelen, Dann Goense, and Koen Langendoen. Radio wave propagation in potato fields. In *1st Workshop on Wireless Network Measurements*, volume 2, pages 331–338. Citeseer, 2005.

[96] Sungsoon Cho and John P Hayes. Impact of mobility on connection in ad hoc networks. In *Wireless Communications and Networking Conference, 2005 IEEE*, volume 3, pages 1650–1656. IEEE, 2005.

[97] Reinhard Diestel. Graph theory (graduate texts in mathematics). 2005.

[98] Béla Bollobás. *Random graphs*. Springer, 1998.

[99] Ji Li, Lachlan LH Andrew, Chuan Heng Foh, Moshe Zukerman, and Hsiao-Hwa Chen. Connectivity, coverage and placement in wireless sensor networks. *Sensors*, 9(10):7664–7693, 2009.

[100] P ERDdS and A R&WI. On random graphs i. *Publ. Math. Debrecen*, 6:290–297, 1959.

[101] Martin Haenggi and Daniele Puccinelli. Routing in ad hoc networks: a case for long hops. *Communications Magazine, IEEE*, 43(10):93–101, 2005.

[102] Uroš M Pešović, Jože J Mohorko, Karl Benkič, and Žarko F Čučej. Single-hop vs. multi-hop–energy efficiency analysis in wireless sensor networks. In *18th Telecommunications Forum, TELFOR*, 2010.

[103] Iordanis Koutsopoulos and Slawomir Stanczak. The impact of transmit rate control on energy-efficient estimation in wireless sensor networks. *Wireless Communications, IEEE Transactions on*, 11(9):3261–3271, 2012.

[104] Baochun Li. End-to-end fair bandwidth allocation in multi-hop wireless ad hoc networks. In *Distributed Computing Systems, 2005. ICDCS 2005. Proceedings. 25th IEEE International Conference on*, pages 471–480. IEEE, 2005.

[105] Rincy Thomas and PK Poonguzhali. An energy efficient multi-sink clustering algorithm for wireless sensors network.

[106] Szymon Fedor and Martin Collier. On the problem of energy efficiency of multi-hop vs one-hop routing in wireless sensor networks. In *Advanced Information Networking and Applications Workshops, 2007, AINAW'07. 21st International Conference on*, volume 2, pages 380–385. IEEE, 2007.

[107] Hui Wang, Nazim Agoulmine, Maode Ma, and Yanliang Jin. Network lifetime optimization in wireless sensor networks. *Selected Areas in Communications, IEEE Journal on*, 28(7):1127–1137, 2010.

[108] P Shrivastava and SB Pokle. A multiple sink repositioning technique to improve the energy efficiency of wireless sensor networks. In *Advanced Computing Technologies (ICACT), 2013 15th International Conference on*, pages 1–4. IEEE, 2013.

[109] Euisin Lee, Soochang Park, Fucai Yu, and S-H Kim. Exploiting mobility for efficient data dissemination in wireless sensor networks. *Communications and Networks, Journal of*, 11(4):337–349, 2009.

[110] Prabhat Singh, Ravi Kumar, and Vinay Kumar. An energy efficient grid based data dissemination routing mechanism to mobile sinks in wireless sensor network. In *Issues and Challenges in Intelligent Computing Techniques (ICICT), 2014 International Conference on*, pages 401–409. IEEE, 2014.

[111] Vahid Shah-Mansouri, A-H Mohsenian-Rad, and Vincent WS Wong. Lexicographically optimal routing for wireless sensor networks with multiple sinks. *Vehicular Technology, IEEE Transactions on*, 58(3):1490–1500, 2009.

[112] Yu Gu, Yusheng Ji, Jie Li, and Baohua Zhao. Eswc: efficient scheduling for the mobile sink in wireless sensor networks with delay constraint. *Parallel and Distributed Systems, IEEE Transactions on*, 24(7):1310–1320, 2013.

[113] E Ilker Oyman and Cem Ersoy. Multiple sink network design problem in large scale wireless sensor networks. In *Communications, 2004 IEEE International Conference on*, volume 6, pages 3663–3667. IEEE, 2004.

[114] Andrej Bogdanov, Elitza Maneva, and Samantha Riesenfeld. Power-aware base station positioning for sensor networks. In *INFOCOM 2004. Twenty-Third AnnualJoint Conference of the IEEE Computer and Communications Societies*, volume 1. IEEE, 2004.

[115] Joakim Flathagen, Øivind Kure, and Paal E Engelstad. Constrained-based multiple sink placement for wireless sensor networks. In *Mobile Adhoc and Sensor Systems (MASS), 2011 IEEE 8th International Conference on*, pages 783–788. IEEE, 2011.

[116] Jayita Barman. *A Study On Sink Mobility In Partitioned Wireless Sensor Network*. PhD thesis, Jadavpur University Kolkata, 2013.

[117] Wint Yi Poe and Jens B. Schmitt. Placing multiple sinks in time-sensitive wireless sensor networks using a genetic algorithm. In *Measuring, Modelling and Evaluation of Computer and Communication Systems (MMB), 2008 14th GI/ITG Conference -*, pages 1–15, March 2008.

[118] Haeyong Kim, Yongho Seok, Nakjung Choi, Yanghee Choi, and Taekyoung Kwon. Optimal multi-sink positioning and energy-efficient routing in wireless sensor networks. In *Information Networking. Convergence in Broadband and Mobile Networking*, pages 264–274. Springer, 2005.

[119] Alain Alcouffe and Gilles Muratet. Optimal location of plants. *Management Science*, 23(3):267–274, 1976.

[120] Umit Akinc and Basheer M Khumawala. An efficient branch and bound algorithm for the capacitated warehouse location problem. *Management Science*, 23(6):585–594, 1977.

[121] Robert R Boorstyn and Howard Frank. Large-scale network topological optimization. *Communications, IEEE Transactions on*, 25(1):29–47, 1977.

[122] Jaap Haartsen. Bluetooth-the universal radio interface for ad hoc, wireless connectivity. *Ericsson review*, 3(1):110–117, 1998.

[123] ZigBee Alliance. Zigbee specification, 2006.

[124] Michel Mouly, Marie-Bernadette Pautet, and Thomas Foreword By-Haug. *The GSM system for mobile communications*. Telecom Publishing, 1992.

[125] Sonali Chouhan, Ranjan Bose, and M Balakrishnan. A framework for energy-consumption-based design space exploration for wireless sensor nodes. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, 28(7):1017–1024, 2009.

[126] Hongbo Jiang, Shudong Jin, and Chonggang Wang. Prediction or not? an energy-efficient framework for clustering-based data collection in wireless sensor networks. *Parallel and Distributed Systems, IEEE Transactions on*, 22(6):1064–1071, 2011.

[127] Moustafa Youssef, Adel Youssef, and Mohamed Younis. Overlapping multihop clustering for wireless sensor networks. *Parallel and Distributed Systems, IEEE Transactions on*, 20(12):1844–1856, 2009.

[128] Okundu Omeni, A Wong, Alison J Burdett, and Christofer Toumazou. Energy efficient medium access protocol for wireless medical body area sensor networks. *Biomedical Circuits and Systems, IEEE Transactions on*, 2(4):251–259, 2008.

[129] Haowen Chan and Adrian Perrig. Ace: An emergent algorithm for highly uniform cluster formation. In *Wireless Sensor Networks*, pages 154–171. Springer, 2004.

[130] Yuhua Liu, Naixue Xiong, Yongfeng Zhao, Athanasios V Vasilakos, Jingju Gao, and Yongcan Jia. Multi-layer clustering routing algorithm for wireless vehicular sensor networks. *IET communications*, 4(7):810–816, 2010.

[131] Suman Banerjee and Samir Khuller. A clustering scheme for hierarchical control in multi-hop wireless networks. In *INFOCOM 2001. Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, volume 2, pages 1028–1037. IEEE, 2001.

[132] Mainak Chatterjee, Sajal K Das, and Damla Turgut. Wca: A weighted clustering algorithm for mobile ad hoc networks. *Cluster Computing*, 5(2):193–204, 2002.

[133] Georgios Smaragdakis, Ibrahim Matta, and Azer Bestavros. Sep: A stable election protocol for clustered heterogeneous wireless sensor networks. Technical report, Boston University Computer Science Department, 2004.

[134] Vivek Mhatre and Catherine Rosenberg. Homogeneous vs heterogeneous clustered sensor networks: a comparative study. In *Communications, 2004 IEEE International Conference on*, volume 6, pages 3646–3651. IEEE, 2004.

[135] Kemal Akkaya and Mohamed Younis. A survey on routing protocols for wireless sensor networks. *Ad hoc networks*, 3(3):325–349, 2005.

[136] Ossama Younis, Marwan Krunz, and Srinivasan Ramasubramanian. Node clustering in wireless sensor networks: recent developments and deployment challenges. *Network, IEEE*, 20(3):20–25, 2006.

[137] Jiang Zhu, Chung-Horng Lung, and Vineet Srivastava. H-dhac: A hybrid clustering protocol for wireless sensor networks. In *Wireless Communications and Mobile Computing Conference (IWCMC), 2013 9th International*, pages 183–188. IEEE, 2013.

[138] Ramesh Rajagopalan and Pramod K Varshney. Data aggregation techniques in sensor networks: A survey. 2006.

[139] Jae-Young Choi, Jun-Hui Lee, and Yeong-Jee Chung. Minimal hop count path routing algorithm for mobile sensor networks. In *Computer and Computational Sciences, 2006. IMSCCS'06. First International Multi-Symposiums on*, volume 2, pages 616–621. IEEE, 2006.

[140] Jun-hu Zhang, Hui Peng, and Tian-tian Yin. Tree-adapting: an adaptive data aggregation method for wireless sensor networks. In *Wireless Communications Networking and Mobile Computing (WiCOM), 2010 6th International Conference on*, pages 1–5. IEEE, 2010.

[141] Fan Bai and Abbas Jamalipour. 3d-dct data aggregation technique for regularly deployed wireless sensor networks. In *Communications, 2008. ICC'08. IEEE International Conference on*, pages 2102–2106. IEEE, 2008.

[142] Tri Pham, Eun Jik Kim, and Melody Moh. On data aggregation quality and energy efficiency of wireless sensor network protocols-extended summary. In *Broadband Networks, 2004. BroadNets 2004. Proceedings. First International Conference on*, pages 730–732. IEEE, 2004.

[143] Ronald R Yager. Intelligent control of the hierarchical agglomerative clustering process. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 30(6):835–845, 2000.

[144] Chung-Horng Lung and Chenjuan Zhou. Using hierarchical agglomerative clustering in wireless sensor networks: An energy-efficient and flexible approach. *Ad Hoc Networks*, 8(3):328–344, 2010.

[145] Margaret H Dunham. *Data mining: Introductory and advanced topics*. Pearson Education India, 2006.

[146] Charles Romesburg. *Cluster analysis for researchers*. Lulu. com, 2004.

[147] Yousef EM Hamouda and C Phillips. Adaptive sampling for energy-efficient collaborative multi-target tracking in wireless sensor networks. *IET wireless sensor systems*, 1(1):15–25, 2011.

[148] J Yan and L Wu. Base station identification in single frequency network positioning system using fuzzy logic technique. *IET wireless sensor systems*, 2(3):230–237, 2012.

[149] Mohamed K Watfa, William Daher, and Hisham Al Azar. A sensor network data aggregation technique. 2009.

[150] Kai-Wei Fan, Sha Liu, and Prasun Sinha. Structure-free data aggregation in sensor networks. *Mobile Computing, IEEE Transactions on*, 6(8):929–942, 2007.

[151] John Heidemann, Fabio Silva, Chalermek Intanagonwiwat, Ramesh Govindan, Deborah Estrin, and Deepak Ganesan. Building efficient wireless sensor networks with low-level naming. In *ACM SIGOPS Operating Systems Review*, volume 35, pages 146–159. ACM, 2001.

[152] Samuel R Madden, Michael J Franklin, Joseph M Hellerstein, and Wei Hong. Tinydb: an acquisitional query processing system for sensor networks. *ACM Transactions on database systems (TODS)*, 30(1):122–173, 2005.

[153] Xu Xu and Weifa Liang. Placing optimal number of sinks in sensor networks for network lifetime maximization. In *Communications (ICC), 2011 IEEE International Conference on*, pages 1–6. IEEE, 2011.

[154] Jiuqiang Xu, Wei Liu, Fenggao Lang, Yuanyuan Zhang, and Chenglong Wang. Distance measurement model based on rssi in wsn. *Wireless Sensor Network*, 2(08):606, 2010.

List of Publications

1. T. Jain, "Wireless environmental monitoring system (WEMS) using data aggregation in a bidirectional hybrid protocol" 6th International Conference, ICISTM 2012, Grenoble, France, March 2012, pp 414-420. [Springer, Book Chapter][Scopus]

2. T. Jain, D. S. Saini and S. V. Bhooshan, "Performance Analysis of Hierarchical Agglomerative Clustering in a Wireless Sensor Network using Quantitative data" 2nd International Conference on Information Systems and Computer Networks (ISCON-2014) , March 1-2, 2015, Mathura, India. [IEEE Conference][Scopus]

3. T. Jain, D. S. Saini and S. V. Bhooshan, "Increasing Lifetime of a Wireless Sensor Network using Multiple Sinks", Proc. In 11th International Conference on Information Technology: New Generations (ITNG 2014), April 7-9, 2014, Las Vegas, Nevada, USA. [IEEE Conference] [Scopus]

4. T. Jain, D. S. Saini and S. V. Bhooshan, "Cluster Head Selection in a Homogeneous Wireless Sensor Network Ensuring Full Connectivity with Minimum Isolated Nodes", Journal of Sensors, Volume 2014, Article ID 724219, 8 pages, doi: 10.1155/2014/724219. [SCI, Scopus]

5. T. Jain, D. S. Saini and S. V. Bhooshan, "Lifetime Optimization of a Multiple Sink Wireless Sensor Network through Energy Balancing", Journal of Sensors, Volume 2015, Article ID 921250, 6 pages. [SCI, Scopus]

List of Other Related Publications

1. A. Chauhan, R. Bhargove, T. Sharma and T. Jain, "A Framework for Energy Efficient Routing Protocol for Homogeneous Wireless Sensor Networks Using Sensing Range" International Conference on Signal Processing & Integrated Networks (SPIN 2014), February 20-21, 2014, Noida, India. [IEEE Conference][Scopus]

2. D. Kashyap, T. Jain, "Comparative analysis of time synchronization schemes for wireless sensor networks". Proceedings of the International Conference on Signal Propagation and Computer Technology(ICSPCT), 12-13 July 2014, pp.106-109, Ajmer, India. [IEEE Conference][Scopus]

3. A. Arman, V. Kher, T. Jain, and D. S. Saini "Implementation of MPLS Routing in Traditional LEACH" International Conference on Signal Processing & Integrated Networks (SPIN 2015), February 17-18, 2015, Noida, India. [IEEE Conference][Scopus]

4. A. Singh, S. Nikhil, S. Rajoriya, and T. Jain, "Design Constraint in Single-hop and Multi-hop Wireless Sensor Network Using Different Network Model Architecture" International Conference on Computing, Communication and Automation (ICCCA 2015), May 15-16, 2015, Greater Noida, India. [IEEE Conference][Scopus]

Bio-Data

1	Name:	Tapan Kumar Jain
2	Date of Birth:	25-Oct-1979
3	Educational Qualifications:	June-2001 B.E. in Electronics and Telecommunications Engineering (72%), Rajiv Gandhi Prodyogiki Vishvavidyalaya, Bhopal, India July-2005 M.Tech in (Electronic Computer) (7.66 CGPA), College of Engineering, Pune (COEP), India May-2015 Ph.D (Thesis submitted), Jaypee University of Information Technology Solan, India
4	Permanent Address:	S/O Vijay Kumar Jain 319, Sanjeevani Nagar, Jabalpur, MP, India Pincode-482003
5	Mobile:	(+91) 9736226189
6	Email:	tapankumarjain@gmail.com