A DISSERTATION
ON

BALANCED CLUSTERING APPROACH TO OPTIMIZE THE LIFETIME OF WIRELESS SENSOR NETWORKS

SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS
FOR THE AWARD OF THE DEGREE
OF

MASTER OF TECHNOLOGY
IN
COMPUTER SCIENCE AND ENGINEERING

Submitted by:

Laksh

University Roll No.:- 2K14/CSE/09

Under the supervision of

Mr. Rajesh Kumar Yadav

Assistant Professor, Computer Science and Engineering Department, DTU

2014-2016

COMPUTER SCIENCE AND ENGINEERING DEPARTMENT

DELHI TECHNOLOGICAL UNIVERSITY

DELHI – 110042, INDIA

ACKNOWLEDGEMENT

First of all, I would like to express my deep sense of respect and gratitude to my project supervisor Mr. Rajesh Kumar Yadav for providing the opportunity of carrying out this project and being the guiding force behind this work. I am deeply indebted to him for the support, advice and encouragement he provided me without which the project could not have been a success.

I am also grateful to Dr. O.P Verma, HOD, Computer Science and Engineering Department, DTU for his immense support. I would also like to acknowledge Delhi Technological University library and staff for providing the right academic resources and environment for this work to be carried out.

Last but not the least I would like to express sincere gratitude to my parents, friends and seniors for constantly encouraging me during the completion of work.

Laksh
University Roll no: 2K14/CSE/09
M.Tech (Computer Science and Engineering)
Department of Computer Science and Engineering
Delhi Technological University
Delhi – 110042

DECLARATION

I hereby declare that the dissertation entitled "*Balanced Clustering Approach to Optimize the Lifetime of Wireless Sensor Network*" which is being submitted to Delhi Technological University, in partial fulfillment of requirements for the award of degree of Master of Technology (Computer Science and Engineering) is a bonafide work carried out by me. The material contained in the report has not been submitted to any university or institution for the award of any degree.

Laksh
2K14/CSE/09

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

CERTIFICATE

This is to certify that the dissertation titled "Balanced Clustering Approach to Optimize the Lifetime of Wireless Sensor Network" is submitted by Laksh, 2K14/CSE/09 in partial fulfillment of the requirements for the award of degree of Masters in Technology in Computer Science & Engineering at Delhi Technological University. It embodies the work carried out by her under my supervision.

Place: Delhi

Date: _____________

Mr. Rajesh Kumar Yadav

SUPERVISOR

Assistant Professor
Department of Computer Science and Engineering
Delhi Technological University

ABSTRACT

Wireless Sensor Network is made up of a huge number of devices known as the sensor nodes. The nodes installed are confined with the battery and other resources. Energy conservation of the nodes is the prime issue in all the applications. As the overall lifetime of the network depends on the energy of these nodes, so maintaining the energy of the nodes becomes an important issue.

This dissertation deals with the network which has nodes whose position is not fixed. The main concern in this work is to balance the load on the all the cluster heads. To balance the load on the CHs, every CH has been confined to the cluster radius; it can form a cluster of only the specified size. This size depends upon the distance of the CH from the base station and its number of neighbors. Those CHs that are close to the BS and have small degree will have a cluster size larger as compared to those that are away from the BS and have large degree.

The selection of CHs is prioritized so as to selects those nodes as the CH which have better parameter values. The priority for each parameter depends upon the criticality in the network. The parameters considered while selecting the CHs are the node degree, mobility, energy and distance from BS.

After comparing the proposed work with the current algorithm, it shows that the lifetime of the network has increased to a desirable amount.

Keywords: Wireless Sensor Networks, mobility, cluster head, energy consumption, node degree, distance, clustering, weight.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVATIONS

ADC	Analog to Digital Convertor
ALERT	Automated Local Evaluation in Real Time
BS	Base Station
CH	Cluster Head
CORIE	Columbian River
DARPA	Defence Advanced Research Projects Agency
DSN	Distributed Sensor Network
EECA	Energy Efficient Clustering Algorithm
B-EECA	Balanced Energy Efficient Clustering Algorithm
EEUC	Energy Efficient Uneven Clustering
EOFS	Environmental Observation and Forecasting System
GDI	Great Duck Island
GPS	Global Positioning System
LEACH	Low Energy Adaptive Clustering Hierarchy
LEACH-M	Mobile Low Energy Adaptive Clustering Hierarchy
MAC	Medium Access Control
MEMS	Micro-Electro-Mechanical System
MTE	Minimum Transmission Energy
OSI	Open System Interconnection
SHM	Structural Health Monitoring
SOSUS	Sound Surveillance System
TCP/IP	Transmission Control Protocol/Internet Protocol
TDMA	Time Division Multiple Access
WSNs	Wireless Sensor Network

CHAPTER 1

INTRODUCTION

1.1 Wireless Sensor Network

Wireless sensor network is the assembly of randomly placed nodes whose function is to sense the environment and provide information to the sink. These nodes collectively perform the desired task and communicate without having dedicated links between them i.e. without wires.

The definition of WSN, according to, SmartDust program of DARPA is: "A sensor network is a collection of a large number of small, battery powered, inexpensive devices which can sense and compute the information and communicate it with other devices in the network to make a global decision about the environment" [1].

1.1.1 Evolution of Sensor Network

During the cold war, United States initiated the development of the sensor networks [1]. A network of auditory devices was placed at different locations on the bottom of the ocean, which could recognize as well as keep a track on Soviet submarines. This system was known as the Sound Surveillance System (SOSUS), because it was used to sense voice. But this network was wired which had no constraint for bandwidth.

In around 1980, the modern research on the sensors began, with the Distributed Sensor Networks (DSN) program at the Defense Advanced Research Projects Agency (DARPA). This includes the sensors (acoustic), algorithms, communication (protocols which are used to link various processes working in a resource-sharing network on a common application), processing techniques and distributed software (dynamically changeable distributed system) [1]. With the advancement in computing and communication, the research in sensor network has rendered and urged it to attain the leading vision. Microelectromechanical system (MEMS) technology based sensors, which are small and inexpensive, wireless networking and processors that are inexpensive and low-powered are used for the deployment of wireless ad-hoc network. Hence, these new techniques are suitable for ad hoc environments which are highly dynamic.

Table 1.1: Three generations of sensor nodes [1]

	1980-1990	2000-2003	2010
Manufacturer	Custom contractors, e.g., for TRSS	Commercial: Crossbow Technology, Inc. Sensoria Corp., Ember Corp.	Dust, Inc. and others to be formed
Size	Large Shoe box and up	Pack of cards to small shoe box	Dust particle
Node Architecture	Separate sensing, processing and communication	Integrated sensing, processing and communication	Integrated sensing, processing and communication
Weight	Kilograms	Grams	Negligible
Topology	Point-to-point, star	Client server, peer to peer	Peer to peer
Deployment	Vehicle-placed or air-drop single sensors	Hand-emplaced	Embedded, "sprinkled" left-behind
Power supply lifetime	Large batteries; hours, days and longer	AA batteries; days to weeks	Solar; months to years

1.1.2 Network Model

Unlike the ad-hoc network, WSN has higher number of nodes, limited resources, and are closely deployed. The topology of the WSN invariably changes and communication is done by broadcast messages. Also the nodes are not identified globally.

The various components of the wireless sensor networks are:

- **Sensor Field:** it is the area in which the devices are deployed.
- **Sensor nodes:** These are the devices that sense the environment and provide the information to the sink.
- **Cluster:** The cluster is formed by grouping of the sensor nodes. This grouping is done based on some parameters for e.g. Energy and distance. Cluster formation helps in effectively utilizing the energy of the devices.
- **Cluster Head:** There is a leader in every cluster thus formed in the network; this leader is known as the cluster head. Each cluster has one CH which has some additional tasks to be preformed. The selection of cluster is done either randomly or using some other parameters. All the non-CH nodes forward their sensed data to the CH for further processing.
- **Sink:** Sink is the node which collects the data from the sensor nodes and processes the data so that it can be used by the user. The sink is used to reduce the overall energy dissipation by reducing the messages sent in total.
- **Task Manager:** Task manager is the Base Station, who is the fundamental point of the network and the control point. The role of the BS is to gather information from the entire network and revert back the control signals into the network. The BS is the gateway between various networks. Any workstation or a laptop can act as a BS.

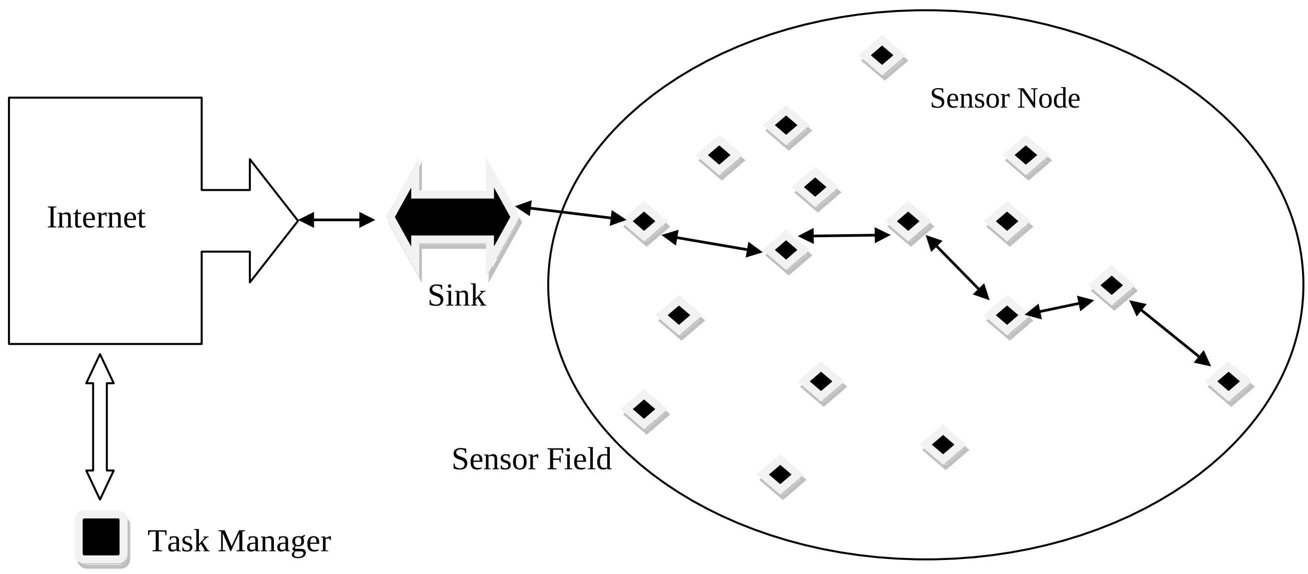

Figure 1.1 Components of Wireless Sensor Networks

1.1.3 Sensor Node

The sensor node is a device, small in size which works on micro-sensor technology, has low processing capability, low battery life, small computation range and small computation power. The basic components of the sensor node are the: sensor, processor, radio transceiver and power supply [4]. The sensor node has various tasks to be done such as gathering environmental information, processing the sensed information, storing and routing the data. The information that is sensed is transformed into electrical signals. **Figure 1.2** portrays the architecture of the sensor node. The major components are:

- **Sensing Unit**: This unit may contain more than one sensor nodes. The nodes are used to sense the data which is then transformed into analog signals. The ADC in the sensing unit is used to convert this analog signal to digital signal.
- **Processing Unit:** The processor in the processing unit is the micro controller unit used to control the processing of the data in the form of digital signals received from the sensing unit. The data is stored in the storage unit of the processing unit.
- **Transceiver:** The role of the transceiver is to connect the network and the node along with controlling the communication between various nodes.
- **Power Unit:** The power or energy given to the sensor unit comes from the power unit. The battery provided here is non-rechargeable and the power is in the form of DC signals.

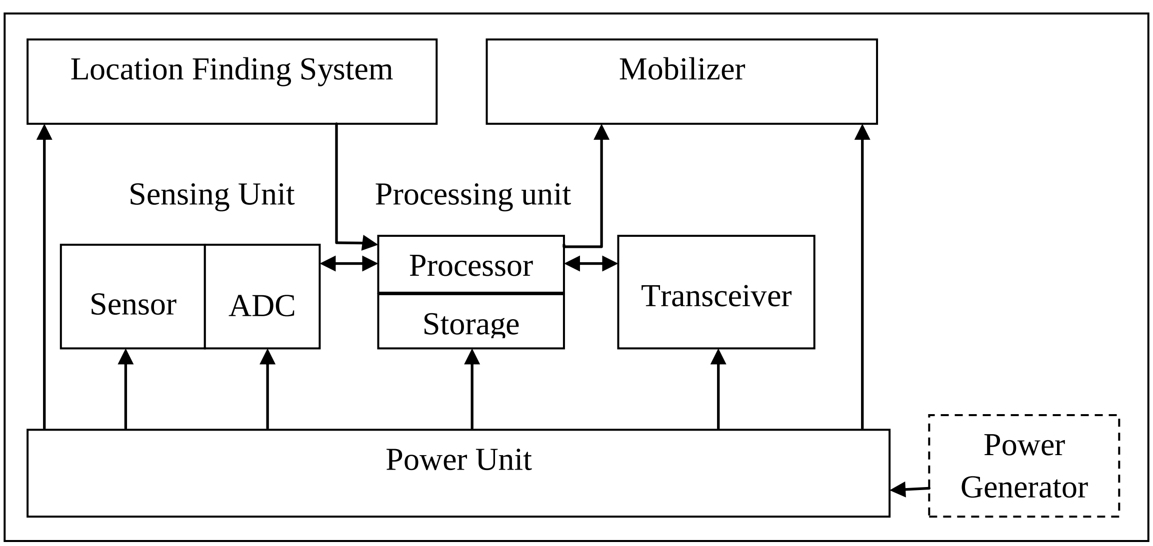

Figure 1.2 Components of a Wireless Sensor Node

Various other components involved are the location finding system, mobilizers which are used in some applications to move the node and the power generator.

1.1.4 WSN Communication Architecture: Protocol Stack

The WSN protocol stack is quite similar to protocol stack of TCP/IP, but there are three planes which take care of the sensor network issues. The protocol stack combines routing awareness and power, network protocols are integrated with data and the wireless medium is used for efficient communication of power. **Figure 1.3** shows the protocol stack.

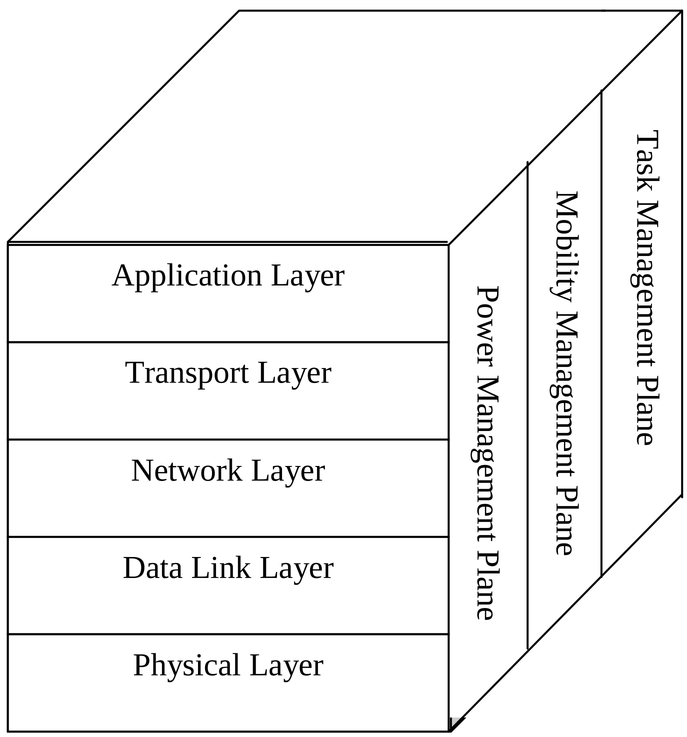

Figure 1.3 Protocol Stack

The different layers of the protocol stack are: the physical layer, the data link layer, the network layer, the transport layer and the application layer. The three planes are: the power management plane, the mobility management plane and the task management plane.

- The Physical Layer has the responsibility for modulation, carrier frequency generation, data encryption, detecting the signal and transmission and receiving mechanism.
- The Data Link Layer is responsible for medium access, multiplexing, error control and lost data frame detection. It ensures reliable point to point and point to multi-hop connection. The data link layer has the MAC layer which does the collision detection.
- The Network Layer is responsible for routing.
- The Transport Layer is responsible for maintaining the flow of the data whenever required. Also it is required when internet is used to access the sensor network.
- The Application Layer presents all the information to the application and the requests are propagated from this layer to the lower layers.
- The Power Management Plane manages the sensor nodes power utilization, due to their battery limitation.
- The Mobility management plane tracks the motion of the sensor nodes to balance the power consumption.
- The Task management plane is responsible for scheduling the sensing and forwarding task of the nodes. It helps in reducing the battery consumption by taking care of the nodes not sensing the data at the same time when not required. All the management planes are designed such that the efficiency of the networks is increased, as the resources are limited in the network.

1.1.5 Characteristics of WSN

The various characteristics of WSN are:

- **Self-organizing**

  The network created by sensor nodes is voluntary and the location of the nodes is random and not predefined. For a self-organizing network, there is no need for it to link to an already established network and the sensor nodes can manage the network communication.

- **Limited communication range and multi-hop routing**

  The sensor nodes have very limited communication range, which is, they can transmit only to a small distance. Multi-hop routing mechanisms are used to overcome this limitation, which consumes less power as compared to single hop communication.

- **Cooperation between sensor nodes**

  Due to the resource limitation, the nodes play varied roles in the network to efficiently transmit data.

- **Dynamic topology**

  Network topology changes from time to time as the nodes die and new nodes join the network.

- **Redundant data**

  As a collection of nodes is deployed in a small region, so many nodes sense the same information and transmit it further. This causes redundant data in the network being transmitted. Due to processing of redundant data, the network efficiency degrades and so the lifetime.

- **Limited resources: energy, storage, processing**

  The wireless sensor network suffers from many resource constraints.

- **Application specific**

  The sensor network is design for a specific kind of task. So for different application, a different network needs to be designed, as it is difficult to modify the deployed network.

1.1.6 Applications of WSNs

Current research in wireless communications has made a practical illusion of affordable, low power deployment of sensor networks. The sensors are used in a number of applications to sense the information according to the application. The sensing can be regular or irregular. The regular sensing activity is performed while sensing the environment like temperature, noise, soil fertility, humidity etc. while the irregular sensing involves sensing of information related to temperature rise of the furnace, detecting border intrusion, stress measurement on machineries. The major applications areas involve the military application, industrial application, environmental monitoring, habitat monitoring, and health application.

- **Military Applications:** Military operation is a wide application area of the sensor network. The characteristics of sensor networks such as low cost, self-organization, dynamic change of topology and fault tolerance are reasonably qualified for battlefield surveillance. A network that supports multicast routing, which is reliable, secured and efficient is essential for military communications, which can be effortlessly satisfied by sensor networks. Sensors are also used in target tracking and target detection. Sensors are used to determine the attack from massive destruction weapons like nuclear, chemical, and biological. Also sensors are helpful in areas where disaster occurs for recovery and peacekeeping.

- **Habitat Monitoring:** The GDI system [2] is one of the applications of habitat monitoring which requires sensor network. The GDI system which collects information about the breeding of the ducks, which is of 7 month duration. As human interference needs to be avoided while monitoring the seabird colonies as it can deform the behavioral patterns, so the sensor nodes play an important role here. After gathering the information it is easy to know what type of nest is better, their incubation occupancy pattern and while the breeding season is on what kind of environmental changes undergo.
  ZebraNet is an application which was used to record the migration of the zebras. The mobility pattern of the location of the zebra is collected by the modus operandi of the application.

- **Health Application:** Sensors nodes are very useful in health applications like managing drugs in the hospital, keeping track of patient or the doctor, controlling the physiological data of the patient, cancer detection, heart rate pointer, organ monitor etc. Evolution of artificial retina [2] is a major project in this area. This project inculcated in the human eye about 100 sensors embedded in a retina prosthesis chip. The patient who had no sight was able to satisfactorily see with this chip. Implanting of sensors in the human body is a disputing issue as the human body should not be vulnerable to any kind of side effects, so the system should be very reliable, safe and such that maintenance required is minimum.

- **Environmental Monitoring:** EOFS is a distributed system that keeps track of the environmental pollution, flooding, etc. [2]. Its extent is over a wide geographical area. WSNs functionality is to overcome and minimize the loss due to disasters. Forest fires are also determined using sensors which are deployed in the forest. The molecules released by the chemical compounds are detected and with the help of GPS, the location of the fire is found. Also the occurrence of fires can be prevented, this can be done by timely monitoring of the temperature range of the area, if there is variable of amount of change in the reading then the nodes forward the information to the BS and corresponding action is taken. CORIE [2] model is set up over the Columbian river which consists of 13 sensor nodes that are fixed and 1 sensor station which is mobile and placed off shore. ALERT [2] is another Wireless Sensor Network which is a real world network deployed in 1970's in California and Arizona, which is used to provide data about the rainfall and the water level. The information provided is useful in detecting the possibility of the flood in that area. This model is prepared using various types of sensor used to sense different kinds of data like water level, hydrological sensor, wind sensor and temperature. The sensed information is transmitted via light to the BS.

- **Structure Health Monitoring System:** WSNs in SHM is used to detect any local damage, damage to the building, measurement of the extent of damage caused and the prediction of the structure durability. To calculate the stability of the multistory buildings or the highways, sensor nodes are installed there. To detect the amount of destruction, the stress and strain values, the temperature value of the various specified points are sensed. The first wireless system was installed in New York to monitor the bridge.

- **Precision Agriculture:** Recent development in WSN has made possible the monitoring of the greenhouse effect [3]. Protection of crops from any kind of damage has become the major issue for the farmers due to varied change in the climatic conditions and the rainfall distribution. The sensors are used to detect the temperature, humidity, light and pressure so that any kind of risk can be reported timely.

- **Industrial Application:** In industrial applications, sensors are used for automation [7] and process control. The various areas where sensor networks are used are real time tracking of the inventory system, inspection of gas leakage and air pressure, real time tracking of machine's health and remote tracking of contaminated areas.

- **Home Appliance and Other Applications:** Various electronic devices act as sensor nodes by common people [2]. The sensors are used in smart phones and smart televisions which function by tracking the motion. Also location sensors, touch sensors and gyrometer sensors are very common and used in many appliances. The education system is also deployed with sensors to enhance the learning and teaching ability.

1.1.7 Challenges and design issues in WSNs

Although WSNs are being used in many applications, but it undergoes various challenges and issues. Various architectures and design constraints are considered depending upon the application, as every application has different requirements [4]. The major necessity lies in escalating the duration of the network by adequately sensing the required information and using energy efficient schemes. The various challenges and issues confronted by WSNs are [5, 6]:

- **Node Deployment:** Node deployment is the positioning of the nodes in the environment. The position of the nodes depends on the application of the WSN as it completely alters the capability of the protocol used. The distribution of the nodes can be either fixed or static. When the distribution of nodes is fixed, each node's position is set prior to network setup, all the nodes are placed according to prior positioning and the routing path is also predefined. While in static deployment the position of the nodes is random, nodes are disseminated unevenly in the network. As the distribution is uneven, so the data routing needs to be efficient to make the network fully connected and operative. Also to limit the energy consumption, the position of the base station is significant. Data sometimes needs to be transmitted to BS in multi hops due to the range of the nodes up to which they can transmit is small.

- **Network dynamics:** Architecture of the network is mostly stationary because only a few applications take into account the mobility of the nodes. But sometimes it becomes necessary to consider the mobility of the nodes or the BS. While the nodes are mobile, increasing the network lifetime route firmness is an important factor because sending data to and from moving nodes becomes difficult.

- **Energy considerations:** In sensor networks, energy is the crucial parameter that is considered. The energy utilization of the sensors needs to be as limited as possible, because the network lifetime prolongs with less energy consumption. These nodes have fixed amount of energy and also are non-rechargeable. The routing protocols used should be such that the route for each node consumes minimum energy, also as compared to single hop, multi-hop routing protocols consume less energy since the transmission distance between nodes and the BS decreases when data travels through intermediate nodes. When a node in the route dies ie. Its energy completely dissipates, then the complete topology of the network changes.

- **Data transmission model:** Depending upon the application, the network models are classified as event driven, continuous or time driven, query driven and hybrid models. The time driven models the sensor nodes transmit the data at regular time intervals. Each node after a fixed time interval switches on/off its transmitter and sensor for data transmission. This type of network is also known as proactive network. The event based and the query based models work based on various events generated, when an event is generated then the data is transmitted. The network is known as a reactive network. The hybrid model is the one which is a blend of both types of networks, proactive and reactive. The hybrid network is a good model for energy consumption.

- **Node diversity:** The nodes in the WSN although are same, but some nodes are given different capabilities like energy, transmission range and data processing. The cluster head selected is amongst the nodes in the network, but it can sense data, transmit data as well as aggregate data while the other nodes can only sense and transmit the data. Also the cluster head has extra memory for storing data.

- **Fault tolerance:** In case when any node dies, then the complete network functionality should not be affected. The topology of the network should be changed accordingly and the data should be rerouted to avoid any kind of lose of information in the network. There can be various reasons due to which the node may fail; they are draining of node energy, congestion in the network or any blockage of node because of some environmental effects.

- **Scalability:** The network may consist of any number of nodes, may be hundreds or thousands in number. Any protocol should be such which can perform in any kind of network, i.e. it should not be affected by the node count in the network. There should also not be a performance issue with any protocol design.

- **Network Connectivity:** All the nodes need to be linked to each other directly or indirectly. Maintaining the association of the network is a major challenge in the WSN as it greatly influences the overall lifetime of the network.

- **Transmission Media:** The communication media in sensor networks is the wireless medium (wireless links). But due to wireless links, the medium has various shortcomings which need to be taken care of, the links have high error rate, signal fading. So to preserve the data from any kind of error, proper protocols need to be used for transmission such as TDMA protocol. Also Bluetooth can be used transmission.

- **Mobility:** Although most of sensor networks are static, but some networks require the nodes to be mobile, i.e. the nodes change their position with change in time. In some applications, the BS can also be mobile. In such a network, routing data becomes a tedious task as the current location of all the nodes need to be tracked and the updated path for data transfer needs to be calculated. The sensing data by nodes can also be either static or dynamic. For reactive network, the event sensing is static and when the data needs to be delivered periodically then the event sensing needs to be dynamic.

- **Data Aggregation [8]:** Data from a massive number of nodes is transmitted to one particular head node for transmission to the BS. But as more than one node senses the same area, so there is redundant data that reaches the node. This redundant data is useless and an overhead for the overall network efficiency. So to avoid any redundancy, data is aggregated at node level. Aggregation removes all the redundant and useless information, and transmits only the required information. This reduces the number of transmission to the BS and hence less energy consumption.

- **Security:** Due to wireless links, sensor network is not secure. Data security is the most important issue in any kind of application, as it can be easily interpreted and tampered. So it becomes the major challenge to secure the data and the network.

- **Quality of Service:** QoS demands that data be transmitted within certain time frame and as soon as it is sensed. Various applications that are real time applications are time critical application. QoS assures reliability as well as usability.

1.2 Motivation

Due to the development of Micro Electro Mechanical Systems, wireless communication and limited power designs, the battery operated sensor nodes have been enabled of small size. A set of nodes is deployed in the area to form a Wireless Sensor Network which is connected through wireless links. All the sensor nodes are formed by the various components that are used for sensing, processing, communicating and for power supply. Since, the sensors have the capability of self-organization and can communicate wirelessly; they can be used in civil, military applications and commercial for climate, habitat monitoring, surveillance, medical observation, disaster management [29] etc. But these sensor nodes face some issues of limited and non-rechargeable battery, limited processing capability, short communication range and memory. Apart from the limited processing capability of the sensor nodes, the nodes work together to analyze the data [5, 34, 25]. These sensor nodes sense the data and the useful information is communicated among each other and the Central node. Data transmission from the sensor nodes to the central node (Base station) is either direct or intermediately via other node. The Base station also known as the sink processes all the information received from the sensors. The BS does not have the constraint of limited power supply.

A lot of research is done in WSNs due to its various applications. But the major area of research lies in reducing the energy consumed by the sensor nodes, hence increasing the lifetime of the WSN. The energy consumption can be decreased by selecting appropriate route from the sender to the receiver; also the node count in the route effects the energy consumption. If a route chosen has less number of nodes and their energy is already depleted, but the route is short can lead to decreased network lifetime, but if the route is longer and has more node count then it can increase the network lifetime. The energy required while transmitting the data is quite large as compared to energy required while receiving the sensed data. So to make the network prolong for a longer duration, energy efficient techniques should be used for routing. To increase the lifetime of the network, various routing algorithms are proposed [21, 22, 23].

Although appropriate routing algorithm is required for data transmission, but along with routing, grouping of sensor nodes also can help in achieving the goal. This sensor node grouping is called Clustering and the groups so formed are called the clusters. Each cluster has a node that is the leader of all other nodes in the cluster, this node is known as the Cluster Head (CH). The role of the CH is to collect information about the environment from all other nodes in its cluster, aggregate the data received and then transmit it to the BS. The selection of CH is done considering various factors like remaining energy of the node, the number of nodes which are closer to the node and the distance from the BS.

Data Aggregation [8, 24, 31] is a technique used by all the CHs to remove redundant and useless data. It eliminates more energy dissipation as inappropriate data is removed, thus leading to less data being transmitted as the number of bits transmitted is also an important factor for energy consumption. Also energy lost during data processing is reduced by data aggregation, although data aggregation itself consumes some energy, but it is much less than the overall energy which is being conserved. The aggregated data is then sent to the BS.

After considering all the factors for energy consumption, an important factor that should be dealt with is the Security [13] of the data being transmitted. The network is prone to various kinds of security thefts. The data can be read, tampered or even lost in the network. So to keep the data intact, some encryption criteria are used.

The clustering algorithms can be divided in three categories: centralized, distributed and hybrid clustering. In centralized clustering the CH is fixed at one node, and the remaining nodes are the cluster members. In distributed clustering the CH node is variable, which is the CH change from one node to another depending upon some parameters. The hybrid clustering is the combination of both centralized and distributed.

Also there are two methods to communicate data to the BS; direct transmission protocol and MTE protocol. In direct transmission protocol, the BS receives data from each CH directly whether it is closer to the BS or far away from the BS. The nodes that are away from the BS dissipate their energy early in comparison to those that are close. But in the MTE, there are various intermediate nodes between the CH and the BS. The CH will send the data to one of the intermediate nodes close to it, which will further transmit it to another node and finally reach the BS. This method is much more efficient than the direct transmission.

1.3 Thesis Outline

The Chapter 1 of the thesis gives the introduction. Further, Chapter 2 describes the various applications of the network, challenges faced by wireless sensor networks and its issues. Also detailed description of the routing and clustering techniques is provided. Chapter 3 provides the details of the protocols studied till date that are related to the proposed work done. Simulation and the results for the algorithms have been defined in chapter 4 and chapter 5 gives the conclusion of the entire thesis.

CHAPTER 2

LITERATURE SURVEY

2.1 Routing in WSN

Routing is a procedure by which a path is established from the source node to the destination node for data transmission. Routing is carried at the network layer of the OSI model. The path formed by routing lets the data travel from the source node via various intermediate nodes and reach the destination. The intermediate nodes that form the path are stored in a table known as the routing table. The routing algorithm is used to determine the path and maintain the table.

Various routing techniques are used which are split into two categories: Protocol operation based and Network structure based routing. Network structure routing includes hierarchical routing, flat routing and location based routing while multipath routing, coherent routing, negotiation routing, QoS routing and query based routing fall under Protocol based routing.

2.2 Clustering in WSN

A lot of sensor nodes are set up in the network, if all nodes forward data to the BS at the same time, it would lead to congestion in the network and also the network lifetime will decrease, firstly because the nodes are deployed at random locations, out of which many nodes' location is far from the BS, so to forward data to BS at a wide range would consume very high energy. Secondly, as the nodes send data they have sensed, so the BS will receive redundant information which creates a huge overhead on the BS. Hence, to avoid such conditions clustering is done over the network. The network is segregated into dissimilar groups based on some criteria; this grouping of nodes is known as clustering. Clustering is done to cut down the energy utilization by the nodes and hence enhance the total network lifetime, as all the nodes need not transmit data to the BS. Earlier the clustering was done using K-means clustering or G-means clustering [9], but there algorithms are not valid for WSNs. Various distributed algorithms are used for WSNs that have less complicacy. Every group of nodes formed is known as a cluster. Every cluster has a node that acts as the leader and is known as the Cluster Head (CH). This CH is selected based on some selection criteria. The CHs communicate with the BS to transmit data. The nodes in the cluster forward their data to the CH of their corresponding cluster, then the CH will aggregate the data collected from all the members of the cluster, this aggregated data is further transmitted to the BS. To avoid overhead on the CHs, the new CHs are selected timely so that load is balanced between all the nodes. The inter-cluster and intra-cluster data transmission can be either done in a single hop or it can be multi hop. Single hop communication allows the CH to send data to BS directly, while in multi hop, the data from one CH reaches the BS via other intermediate CHs. Multi hop transmission is done when a CH is located at a distance far away from the BS, so sending data via other nodes is more adequate than direct transmission. Also the nodes have a particular transmission range; only within this distance the nodes can transmit the data. Along with the benefits of clustering, there are some shortcomings too like some extra work load in selecting the CH, forming the cluster and CH allocation to each cluster.

Clustering Process: The process of clustering is done in two steps: the selection of the Cluster Head and formation of the cluster.

To choose the CH, any of the following method can be employed: first, BS centralization in which cluster selection is done by the BS depending on its location from the BS and it remaining energy. Second, Cluster head selection is done by the nodes themselves; the nodes choose the CH among themselves by generating a random number. Third, a hybrid of the two methods can be done [13].

The important aspects in selection of CHs are:

- The distance between the BS and the CH should be such that it’s not too far away from the BS. More distance from the BS leads to high energy consumption which is undesirable.
- The CHs should be equally divided in the network so that there is not much distance from the CHs to its members.
- As all the nodes have same energy, but the task performed by the CHs is more energy consuming so the CHs should be selected again. This re-selection should be done in such a manner so as to maintain equivalence between the nodes’ energy.
- The remaining energy of the nodes should be taken into account while selecting the CHs.

- While selecting the CHs and forming the clusters, there is some time lapse. This time lapse should be taken care of.

The important aspects while forming a cluster are:

- The distance is the most important factor while joining a cluster.
- The size of the cluster is another major issue, depending on the cluster size deciding whether to join a cluster or not is very critical. A large cluster size leads to a huge overload on the CH.
- The hop count from the node to the CH.

The different types of clustering methodologies are described below:

2.2.1 Hierarchical Clustering

In this type of routing, the different clusters are organized into various layers. The lower layer transmits the data to the upper layer and further to the upper layer till the data reaches the BS [11]. It transmits data to the BS in multi hop which decreases the utilization of energy. The data is first aggregated and then forwarded to the next level; this minimizes the data count which is transmitted over the network. The clusters formed are based on the energy that is left for each node. The best suitable example for this type of clustering is the LEACH protocol.

2.2.2 Data Centric Clustering

This clustering method calculates the minimum distance between all pair of sensor nodes [12]; the data is aggregated and further sent to a destination from various resources which use the names data. As the network contains enormous number of nodes, so to differentiate each node by a particular identity becomes difficult. Hence, to identify each node its location, ability and closeness are considered. Considering the performance metric, data-centric approach provides a better performance.

Table 2.1 Comparison of various Clustering methodologies

	Data Centric	**Hierarchical**
Scalability	Limited	Good
Data Diffusion	No	Yes
Lifetime	Long	Long
Power Required	Limited	High

2.3 Related Protocols

A lot of research is being done to find clustering techniques that are energy efficient. All the research revolves around finding the efficient techniques under every category. The various algorithms have been discussed below to understand the work done till current time and the proposed work.

2.3.1 LEACH

Low Energy Adaptive Clustering Hierarchy (LEACH) protocol is the basis for all other protocols proposed for WSN. LEACH is used to lessen the energy drained by the nodes in transmitting data to BS. LEACH is an adaptive clustering protocol, which is self-organizing algorithm. In LEACH the energy load is distributed randomly to evenly divide it in all the sensors. All the nodes at the start have the same energy and other parameters. The nodes themselves form the clusters and random selection of cluster head is done for all the clusters. The nodes in a cluster forward their data to the CH of that particular cluster. All the CHs will aggregate the data collected from nodes along with their own data and then forward the aggregated data to the BS. After a time interruption the CHs are again randomly selected and the whole process is repeated.

LEACH protocol has two functional phases, which are the setup phase and the steady state phase.

Set Up Phase: The clustering and the CH selection is done in the first phase. Every node makes an opinion about whether or not to become a CH for the respective round. This selection is done by a probability function, a threshold value *T(i)* is pre-decided and is correlated with the random number that is generated by each node. This random number has the value between 0 and 1. The value if lower than the threshold, then the node becomes the CH.

$$T(i) = \begin{cases} \frac{P_{ch}}{1-P_{ch} \times (r \bmod \frac{1}{P_{ch}})} & if\ i \in G \\ 0 & otherwise \end{cases} \quad (2.1)$$

Where,

P_{ch} is the CH probability

r is the current round

G is the node set which have not yet been elected as the CH in the previous $1/P_{ch}$ rounds.

Steady State Phase: This phase is the data transmission phase, in which the member nodes forward their data to the CH, and is forwarded to the BS for processing. For data transmission, a slot is appointed to each node in which the nodes forward their sensed data to the CH and later the nodes go to sleep mode until next slot is assigned. This is done to preserve the energy as the nodes while sleeping will not utilize any energy. Then on receiving data of all the members, the CH aggregates the data collected and transmits it to the BS.

There are various advantages of the LEACH protocol. They are:

- All the nodes have equal probability of becoming the CH with a constraint that the same node cannot become the CH for two consecutive rounds. This helps balancing the load on one node.
- As data transmission is done using TDMA protocol, this prevents any kind of collision to occur at the CH.
- Due to aggregation at the CH, traffic is reduced to a high extent in the whole network.

Although there are many advantages of the LEACH protocol, but many disadvantages too exist.

These are:

- As the data transmission is done in single hop, so it cannot be used for large networks as energy in transmitting data by nodes located at a large distance from BS would consumed a huge amount of energy, thus leading to overall decay in network lifetime.
- The network load is not balanced as the CH selection is done depending on the probability.
- The network is not evenly divided in to clusters because the CHs are selected randomly. This results in some CHs having more cluster members while others having very few nodes in the cluster.
- After cluster formation if the CH dies, the cluster becomes useless as the nodes' data would not be received by any CH and also the data would not reach the BS.

2.3.2 Mobile LEACH

Wireless Sensor Networks are used in many applications; some applications require nodes to be static but some real time applications require mobile nodes; the nodes that change their position with time. According to researchers, when the nodes are mobile then the network coverage improves to a large extent. But mobility in the network leads to varied challenges like maintaining the current location of the nodes, data management, routing, network coverage and software support. Apart from these challenges, the most important constraint is to manage the route between the nodes while they are moving, as the location of the nodes keeps changing so to maintain the routing table becomes a major issue at all levels. Also maintaining the cluster information becomes an important issue. When the network is dynamic, the conventional protocols applicable for static network cannot be used, as the complete topology of the network transforms. These protocols need to be optimized according to the network requirement, which requires the study of the nodes' mobility patterns and mobility metrics.

Similar to LEACH, LEACH-M [15] also has two phases of operation: the set-up phase and the steady-state phase. In the set-up phase, the CH is selected. This selection is based on the mobility of the nodes, the nodes that are either static or those which have minimum mobility are only selected as the CHs. In the steady-state phase, every sensor node receives a message according to TDMA schedule after a fixed time interval, thereafter receiving this message the cluster is reformed considering minimum energy consumption.

Although the CH selection in LEACH-M is same as that in LEACH protocol, but the CHs are not fixed and not pre-assigned. In such a case, the CH nodes will be overloaded and the entire network performance would de degraded. To avoid such a condition, the CHs are selected randomly in every round depending on the energy of the nodes. This permits the energy distribution among all the nodes and the energy of a particular node does not drain completely. After the CHs are selected, all the CHs broadcast their status to all other non-cluster head nodes. The normal nodes determine which CH to join depending upon the minimum energy consumption in transmitting data to the CH. After cluster formation, a time schedule is set by the CH for its cluster members, this time schedule is used to let the nodes transmit their data and at the remaining time the nodes go to sleep mode to avoid any kind of unnecessary energy dissipation. But when the CHs are selected depending upon the single energy parameter in mobile network, it leads to severe problems, as the node selected as the CH may have very high mobility and may leave the cluster after some time. At this time, the cluster would become unconnected with the rest of the environment and the nodes would be unable transmit their data. So to resolve this problem, LEACH-M considers the node mobility in selecting the CHs. The nodes that have zero or minimum mobility are only selected as the CHs.

Some important aspects that are considered in LEACH-M are:

- Role: This parameter is used to keep a record of which node is a CH, for CH nodes a value parameter is set to 1 (value=1) and for non-CH nodes it is set to 0 (value=0).
- Cluster members: A CH node has a record of reference to all the nodes that are members of the cluster.
- Mobility: Mobility of a node is computed depending upon either of the two factors: the number of times the nodes changes its cluster or on the basis of remoteness.
- TDMA Schedule: This time slot is maintained to let the CH node know when to collect the data from the cluster members.

All the four factors' information is required to be maintained by the nodes which is required while selection of the CH, but mobility is the most important factor amongst all these. To compute the mobility of the modes there are two distinct approaches. The first method is to calculate it using remoteness [17] and the other method is by finding the number of clusters the node changes. LEACH-M basically uses the approach of remoteness to calculate the mobility of the nodes.

I. Remoteness technique to calculate Mobility

Remoteness is basically the amount of the mobility factor.

Let $u = 0, 1, 2, 3 \ldots Nr-1$, where Nr is the number of nodes

$n_u(t)$ is the location vector of the node u, at any time t

$d_{uv}(t) = | n_v(t) - n_u(t) |$ is the distance between nodes u and v at time t

The remoteness of node u to v is given by

$R_{uv}(t) = Fn(d_{uv}(t))$, where Fn is remoteness function.

In other words, remoteness is the distance between two nodes.

When the nodes' motion is relative to some other nodes, then the remoteness parameters remains same as the earlier value. But when the node movement deviates at an angle and the speed of the node changes, then the remoteness of the node changes. In such a case, the remoteness is given by,

$$M_u(t) = \frac{1}{Nr-1} \sum_{v=0}^{Nr-1} | d'uv(t) | \quad (2.2)$$

II. Transition Count

When the node is mobile, it may leave its cluster and may link with some other cluster in the network. This can happen not only once but many times in the complete lifetime of the node. The mobility of the node is computed using this transition of the node from one cluster to other. The count of the node changing its cluster gives its mobility parameter.

Mobile LEACH has a limitation in selecting the CH. It considers just the mobility of the nodes while selecting the CH, but only taking the mobility would not result in efficient selection of the CHs.

2.3.3 EEUC

Energy Efficient Unequal Clustering (EEUC) protocol [18] is used for periodic information collecting applications. In EEUC, the network is divided into clusters that are of unequal sizes and the data transmission is done through multi hop rather than single hop transmission. The CHs are chosen locally by the nodes unlike the LEACH protocol. The approach of unequal clustering is fulfilled by the radius of each cluster, as the cluster distance increases from the BS so does the radius of the cluster increases. The clusters that are close to the BS are having small cluster size while those away from the BS have clusters of comparatively large size. This size of the clusters is maintained so as to balance the energy utilization for the inter-cluster communication and increment the lifetime of the network. The routing here used is multi hop routing, out of the adjacent nodes of the CH, it selects a relay node depending upon its distance from the BS and energy. The CHs that are near to the BS are used as the routers for the nodes that are located at a wide distance from the BS as it might be the case that the transmission range of the CH away is not up to the BS, so it requires an intermediate router to route its data. Due to this reason, the nodes that lie in the transmitting range of BS and that are at a closer distance to BS would need to transmit other CHs' data along with their own data, leading to high load on these nodes. This inter-cluster communication can lead to decrease in lifetime of the network as the CHs close to BS would die very early due to much load. To avoid this situation, the concept of unequal clustering is used, which states that the size of cluster close to BS is small as it would preserve energy in intra-cluster communication to be used to inter-cluster communication. Due to small cluster size, the CH will have a load of fewer numbers of nodes.

The CHs in each round are elected again so as to disseminate the load amidst all the nodes, and not let only a particular node handle the entire load. The selection of CH is done based on its remaining energy in the network.

The cluster radius [18, 26] of the CH is determined by using the formula 2.3

$$S_p.R_{comp}=\left[1 - c\ \frac{dBSmax - d(sp,BS)}{dBSmax - dBSmin}\right] R^0_{comp} \tag{2.3}$$

Where, dBS_{max} and dBS_{min} are the maximum and minimum distance from the BS

$d(s_p, BS)$ is the distance between node *p* and the BS c is a constant such that $0 < c < 1$.

R^0_{comp} is the predefined radius of the cluster.

Hence, considering the radius of each clustering by this formula, the energy consumed by the nodes has been decreased to a great extent.

The limitation of EEUC algorithm is that it considers just the distance between the nodes and the BS.

2.3.4 EECA

EECA is Energy Efficient Clustering Algorithm [14], this algorithm deals with fixed WSN as well as mobile WSN. The algorithm has proposed three solutions: first, for fixed network, second, for mobile network when the nodes' mobility is static and third, for mobile network when the mobility of the nodes varies with time.

The algorithm looks at the four parameters of the nodes to select the CHs [14, 30]; these are the degree of the node in its transmission range, energy consumed in calculating the degree, distance from the BS and the mobility of the nodes. The weight of each node is calculated combining all the four parameters and the nodes that have the minimum weight are selected as the CHs for the ongoing round. The same process for the CH selection is again repeated and new CHs are selected. According to this method of CH selection, the energy of the nodes is consumed efficiently and the lifetime increased to a great extent.

But the CHs that are selected are those that are close to BS, have high degree, low consumed energy and least mobility. But as the nodes that are close to BS and have high degree have more probability of becoming the CH, in that case the node will have high load as the degree is high and has many nodes in its cluster. Thus the load on the CHs in not balanced and so the energy of the nodes would be consumed very fast and the nodes would die early.

2.4 Research Objective

Various current protocols have been discussed in the previous sections and we learn that there are some limitations faced in each of the algorithm. Those limitations need to be removed as it directly has an impact on the network performance. So to overcome these limitations we carry out this research work with an objective of developing an energy efficient protocol with load balanced energy consumption among sensors to route the data in the wireless sensor networks. The objective of energy efficiency translates into prolonging the network lifetime whereas load balanced energy consumption can be realized by delaying first node death. These objectives depend upon sub objectives such as better cluster formation and cluster head selection techniques.

CHAPTER 3

PROPOSED PROTOCOL

Here we introduce the optimized protocols for homogeneous mobile network. The protocol EECA-M2 [14] described above is optimized to lengthen the lifetime of the network. The proposed protocol deals with the same CH selection process as in EECA-M2, but the parameters that are used for CH selection are each given a priority [27, 28] so as to prioritize which parameter to be considered as utmost importance than other parameters. The node is elected as the CH based on the four parameters; its degree within its transmission range, distance from the BS, mobility and the energy consumed in finding the 1-hop neighbors. Also while forming a cluster; each CH can form a cluster of a defined size, the radius for which depends upon the distance between node and BS as described in EEUC [18] and another proposed work considers various node parameters for defining the radius. Considering the above factors, the protocol increases the network lifetime and balances the load on the CHs.

Below we describe the protocols in detail. First, the complete description of the system model is given and various assumptions considered are defined. Further the detailed description of the proposed protocols is defined and the performance evaluation is done later.

3.1 System Model

The system model constitutes of two models: the energy model and the network model. The energy model explains the energy parameters used by the nodes and the energy usage of the nodes while communicating with other nodes in the network. The network model depicts the network domain and the properties of the sensor nodes.

3.1.1 Network Model

The sensor network is modeled by a graph $Gr = (V, E)$ where V is the vertex of the graph, here each sensor node acts as a vertex and E is the edge which is the link between two nodes. E is represented as

$$E= \{(p,q) \in V / D(p,q) <=R\} \quad (3.1)$$

Where $D(p,q)$ is the Euclidean distance between the two nodes p and q and R is the transmission range of each node. The node count is represented by Nr and has a sink node or the BS. The position of the nodes is unplanned and is disseminated throughout the network. The network size is specified by the coordinate axis having the area Z as $x_c * y_c$, where x_c and y_c are the dimension of the x-axis and y-axis. Various assumption and specifications of the network are detailed below:

- The location where BS is placed is the center of the field; the BS has no limit on power and resources.
- At the set up time, same energy and resources are allocated to the sensor nodes.
- The sensor nodes are not fixed at a particular location, but are mobile and change their position after some time intervals.
- Each node has a particular node id, and the capabilities of the nodes like sensing, communicating and transmitting are same for all.
- Initially the transmission range for all the nodes is same and predefined.
- The nodes sense the information from the environment at a regular basis, but transmission of the information to the CH is done only at a defined time and not at any time.
- Only the CHs interact with the BS and carry the data to the BS.
- The BS collects the data from the CHs, processes it and equips the user with the information required.

3.1.2 Energy Model

The energy model used is the First Order Radio Model used for various protocols for WSN. The model comprises of basic three components: amplifier, transmitter and receiver. The model is shown in **figure 3.1.**

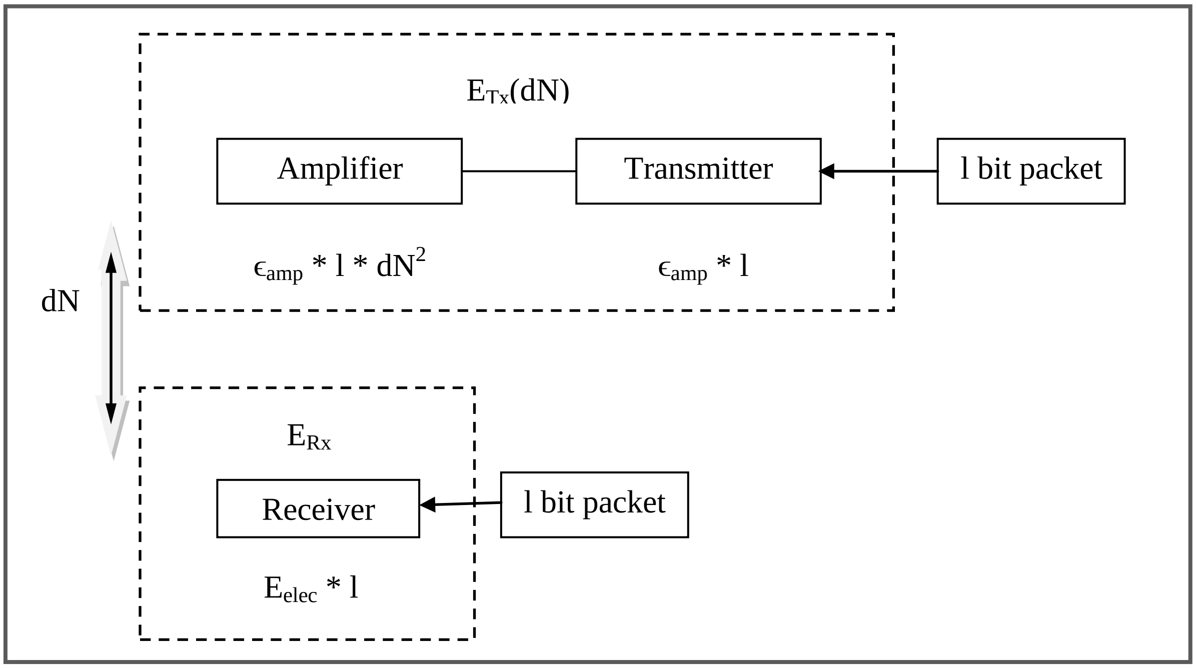

Figure 3.1 First Order Radio Energy Model

The energy model has two different propagation models, the free space and two ray ground propagation model. The difference in these both the models is in regards to the path of the line of sight between the receiver and the sender. The free space model has a direct path between the sender and the receiver while in the two ray ground model, the transmission path between the sensor and the receiver is not direct but has multiple paths.

The functionality of the amplifier is to amplify the received signals as the signal gets week in its path. To measure the energy depletion of the nodes when transmitting the data packet of size 'l' bits to other node situated at a distance 'dN' apart from it, we use the formula 3.2

$$E_{Tx}(l, dN) = \begin{cases} l * E_{elec} \ + \ l * \epsilon_{fs} * dN^2 & , dN < do \\ l * E_{elec} \ + \ l * \epsilon_{mp} * dN^4 & dN \geq do \end{cases} \tag{3.2}$$

Here, E_{Tx} is the energy depleted by the node while transmission of data. The energy consumed directly depends upon the size of the data packet and the distance from the receiving node.

E_{elec} is the electrical energy required for the transmitter to function perfectly for every data unit.

ϵ_{fs} and ϵ_{mp} are two amplification parameters that are used for the different propagation models. ϵ_{fs} is used in free space propagation model and ϵ_{mp} is used in Two ray ground propagation model.

d_o represents the optimal distance and is represented in equation 3.3

$$d_o = \sqrt{\frac{\epsilon_{fs}}{\epsilon_{mp}}} \tag{3.3}$$

The distance is the deciding factor between the two propagation models, if the distance between two nodes leads d_o then two ray propagation model is used else the free space model is used.

Energy utilized at the time when the nodes receives the data is the reception energy E_{Rx} and is given by the equation 3.4

$$E_{Rx}(l) = l * E_{elec} \tag{3.4}$$

3.2 Proposed 1

3.2.1 Problem Statement

The core area in this work is the fixed and mobile network where the nodes are mobile in mobile sensor network. The network model for the static model is illustrated in the previous topic. Apart from the above model, it is that the nodes are dynamic in our algorithm. The problem that we look upon in this work is described below:

The CH selection approach in the energy efficient clustering algorithm used is by computing the weight of individual node, the weight is calculated using the number of neighboring nodes, distance from the BS, remaining energy and its mobility. Those nodes are chosen for the CHs which have minimum weight. So the nodes that are close to the BS and those that have high degree have a high probability of becoming the CH. In such a case, due to large degree the load on the CHs becomes too high. Also the CHs formed will be only close to the BS and not equally distributed in the network.

As a solution to the above problem, we proposed the algorithm described in the next section.

3.2.2 Proposed Solution

This section describes our work proposed1 to optimize the above problem. The protocol is proposed for homogeneous network where each node has equivalent initial energy when the nodes are deployed in the network. The selection of cluster head criteria is based on the following parameters: degree, node and base station distance, remaining energy and mobility. An overall weight for all the nodes is calculated while assigning a priority to each parameter. The nodes with the minimum weight become the CHs. The number of CHs depends upon the network parameters which is the optimal number. After the CHs are selected, each cluster forms a cluster whose radius is determined considering its distance from the BS. The cluster members transmit their data to their CH after which the CH aggregates the data and further transmits it to the BS for further processing. The detailed protocol is explained in the following sections.

3.2.2.1 Features

Some features of the proposed protocol are discussed below. They are:

- Count of cluster heads based on the network dimensions and the number of nodes in the network.
- Introduction of priority for each factor while selecting the CH, a node is selected as the CH if it has the desired properties.
- The selection of the CH depends upon its distance from the BS, degree, current energy and its mobility.
- The nodes with high degree and close to the BS are given more priority.
- Concept of cluster radius is introduced in this protocol.
- The cluster radius is dependent upon the CH’s distance from the BS; more the distance from the BS larger is the radius.

3.2.2.2 Cluster Head Selection Process for Mobile Nodes

The network is partitioned into various clusters and each cluster has a leader node known as the Cluster Head. This cluster head is responsible for all the data processing of its member nodes. All the CHs have extra responsibilities as compared to the non-cluster head nodes. This extra burden consumes enough energy and hence reduces the lifetime of the node. So to avoid a huge load on some particular nodes, the CHs are swapped in each round so that the CH overload is distributes amongst all the nodes and maintains the lifetime of the network. So to choose a CH which consumes least energy and which is optimal for current round becomes a crucial task.

In proposed1 protocol, the selection of CH is done based on the prioritization of various factors of the nodes. Each node has the knowledge about its number of neighboring nodes, node and base station distance and its energy remaining. The mobility for each node is calculated using the transition count method. Transition count method states that the mobility of a node is the number of times it moves from one cluster and joins another cluster in a particular time period. So considering these four factors we find which nodes are suitable to become the CHs.

Stepwise selection of cluster head:

- Every node $p \epsilon V$ transmits a 'hello' message to other nodes to find 1-hop neighbors.
- Every node computes its degree. The degree of each nodes depends upon its number of neighbors, and is given by

  $$Degree\ (p) = |\ E1\ | \tag{3.5}$$

  $$E1 = \{(p,q)\ \epsilon\ V/\ D(p,q) <= R\} \tag{3.6}$$

  R is the transmission range of every node which is predefined.
- Find the energy consumed $E_u(p)$ for all nodes in calculating the degree
- For every node, the distance between BS and p is computed. The distance is calculated as:

  $$Ds\_BS(p) = \sqrt{(x1 - X_{BS})^2 + (y1 - Y_{BS})^2} \tag{3.7}$$

  here, x1 and y1 are the coordinates of u
- For every node, compute the mobility *Mob (p)*. The average mobility [32, 33] of $p \epsilon V$ until present time T is given by:

$$Mob\ (p) = \frac{1}{T}\sum_{ti=1}^{T}\sqrt{(x1 - x2)^2 + (y1 - y2)^2} \tag{3.8}$$

Here, the coordinates of p at time t_i and $(t_i\text{-}1)$ are given by *(x1, y1)* and *(x2, y2)*.

- Now, every node computes its combined weight considering all the parameters above, using the equation 3.9:

$$Wt(p)=\frac{1*w1}{Degree\ (p)}+\frac{w2*E_u(p)}{EI}+\frac{w3*Ds\_BS(p)}{Ds\_MaxBS}+\frac{w4*Mob\ (p)}{Mob_{Max}} \quad (3.9)$$

Where, w1, w2, w3, w4 are the coefficients for the weight parameters [20]. These coefficients are used to give priority to each parameter.

The priority is assigned in such a way such that

$$w1 + w2 + w3 + w4 = 1 \quad (3.10)$$

The values are selected according to the requirement of the protocol. Here we choose the nodes having the minimum weight as the CH. So accordingly we set the values for coefficients, if we want to choose a node that has high degree then we set the coefficient of the degree as a small value, because it would further decrease the degree parameter and the chance of the node to become CH increases. Similarly all the coefficients are set for each factor. In our work we have used the priority coefficients as:

w1=0.18, w2=0.14, w3= 0.15, w4= 0.53

EI is the initial energy of each node

Ds_MaxBS is the maximum distance of any node from the BS

Mob_{max} is the maximum mobility of any node

- The CH is chosen which has the minimum weight. Considering minimum weight, the nodes which have high degree, less consumed energy, close to the BS and has least mobility is selected.
- Repeat complete procedure for the sensors which are remaining and have not yet been assigned as CH.

In each round the number of cluster heads [19] selected is decided by using the formula 3.11

$$N_{CH} = \sqrt{\frac{N_r}{2\pi} * \frac{\epsilon_{fs}}{\epsilon_{mp}}} * \frac{Z}{(Ds_{BSavg})^2} \tag{3.11}$$

Where, d_{BS} is given by equation 3.12:

$$Ds_{BS}avg = \frac{1}{N}\sum_{i=1}^{N_r} d(i, BS) \tag{3.12}$$

$Ds_{BS}avg$ gives the average distance between the BS and all nodes; Z is the network size, Nr the number of nodes in the network.

3.2.2.3 Cluster Head Selection Process for Fixed Nodes

When the network is static and the nodes are fixed i.e. they have fixed locations although the network, the CH selection depends only on three parameters: degree, energy utilized and BS and node distance. According to the values of these three parameters, the weight for each node is calculated and the nodes with the minimum weight are selected as the CHs for the ongoing round. The CHs selected are then swapped for the next round depending upon the new values for each parameter.

In this protocol for fixed nodes we do not assign a priority to the parameters as all the three parameters are as essential as the other. So considering these three factors we find which nodes are suitable to become the CHs as described below.

Stepwise selection of cluster head:

- Every node $p \epsilon V$ transmits a 'hello' message to other nodes to find 1-hop neighbors.
- Every node computes its degree. The degree of each nodes depends upon its number of neighbors, and is given by

$$Degree\ (p) = |\ E1\ | \tag{3.13}$$

$$E1 = \{(p,q) \epsilon V / D(p,q) <= R\} \tag{3.14}$$

  R is the transmission range of every node which is predefined.
- Find the energy consumed $E_u(p)$ for all nodes in calculating the degree
- For every node, the distance between BS and p is computed. The distance is calculated as:

$$Ds\_BS(p) = \sqrt{(x1 - X_{BS})^2 + (y1 - Y_{BS})^2} \quad (3.15)$$

here, x1 and y1 are the coordinates of u

- Now, every node computes its combined weight considering all the parameters above, using the equation 3.16:

$$Wt(p) = \frac{1 * w1}{Degree\ (p)} + \frac{w2 * E_u(p)}{EI} + \frac{w3 * Ds\_BS(p)}{Ds\_MaxBS} \quad (3.16)$$

Where, w1, w2, w3 are the coefficients for the weight parameters [20]. These coefficients are used to give priority to each parameter.

The priority is assigned in such a way such that

$$w1 + w2 + w3 = 1 \quad (3.17)$$

here the coefficient values are kept same for each factor as all the three factors play the same role in the CH selection.

w1= 0.33, w2=0.33, w3=0.33

EI is the initial energy of each node

Ds_MaxBS is the maximum distance of any node from the BS

- The CH is chosen which has the minimum weight. Considering minimum weight, the nodes which have high degree, less consumed energy, and close to the BS is selected.
- Repeat complete procedure for the sensors which are remaining and have not yet been assigned as CH.

In each round the number of cluster heads [19] selected is decided by using the formula 3.18

$$N_{CH} = \sqrt{\frac{Nr}{2\pi} * \frac{\epsilon_{fs}}{\epsilon_{mp}}} * \frac{Z}{(Ds_{BSavg})^2} \quad (3.18)$$

Where, d_{BS} is given by equation 3.19:

$$Ds_{BS}avg = \frac{1}{N}\sum_{i=1}^{Nr} d(i, BS) \qquad (3.19)$$

$Ds_{BS}avg$ gives the average distance between the BS and all nodes; Z is the network size, Nr the number of nodes in the network.

3.2.2.4 Cluster Formation Process

After CHs are selected, all the nodes that are not the CHs make a decision about which CHs would be their leader. In the prevailing protocols [21, 22], the CH for a node is selected based on its distance from the CH, a CH which is closest to the node becomes its head node. But this method is not much efficient as in this method the CH that the node joins may be very far away from the BS which will leads to excess energy consumption, hence a huge lose in energy.

Another protocol [23] proposed a new method for cluster formation that was based on the distance between CH and the BS. Instead of considering the distance between the node and the BS, this method considers the distance between the CH and the BS. But since the energy of the nodes is not considered, so it did not prove to be an efficient method.

But as an improvement to above discussed methods, a new method for cluster formation is proposed. The proposed method defines a cluster radius for each cluster, this radius sets the range for each CH in which is can form a cluster. In this work the cluster's radius is based on the CH's distance from the BS as described in equation 2.3. The formula used is as shown in equation 3.20

$$S_p.R_{comp} = \left[1 - c\ \frac{dBSmax - d(sp, BS)}{dBSmax - dBSmin}\right] R_{comp}^{0} \qquad (3.20)$$

According to this formula, the CH that is close to the BS has a cluster radius of small size while that away from the CH has a radius of larger size. This is done so as to preserve the energy of the CHs that are close to the BS so that they can function when the other nodes in the network are dead and thus increasing the time for which the complete network functions.

Since the distance is considered while deciding the cluster size, so it works out to be an efficient method for cluster formation.

But while cluster formation, many nodes fall under more than one CH, this is so because the radius of two adjacent CHs overlaps each other's radius. To avoid the overlapping of the nodes, the node should join a CH that transmits the data while consuming minimum energy. In I-EECA this is achieved by using the formula 3.21

$$\frac{Dis_{CH\_BS} + Dis_{CH\_N}}{E_{CH}} \tag{3.21}$$

Here, Dis_{CH_BS} is the distance between the BS and the CH

Dis_{CH_N} is the distance between the CH and the node

E_{CH} is the remaining energy of the CH

The node joins the CH which has smaller value for the above formula. This formula considers the energy of the CH and the distances of the CH from the node and the BS. So, the node joins a cluster that has small data transmission distance and more energy. Small data transmission energy refers to the total distance of the CH and the node and the CH and the BS, small the distance less will be the energy consumed. Also energy for the CHs is considered. This phenomenon is explained with an example in **figure3.2.**

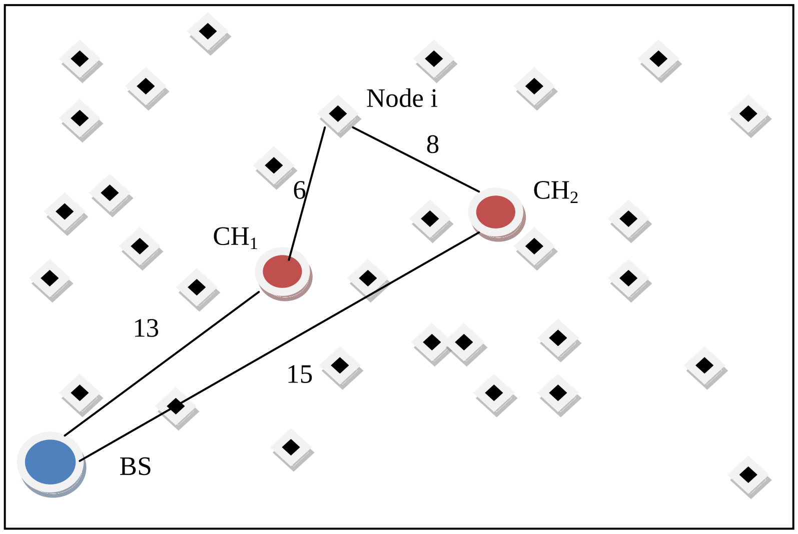

Figure3.2. Network Deployment

The distance between node i and CH_1 is 6 and between CH_1 and BS is 13. Distance between node i and CH_2 is 8 and between CH_2 and BS is 15. The remaining energy of CH_1 is 0.4J and that of

CH_2 is 0.5J. So according to the above formula, the values for CH_1 and CH_2 are 47.5 and 46 respectively. So the node i will join the CH_2 as it gives more efficient value.

3.3 Proposed 2

3.3.1 Problem Statement

As discussed in the proposed1, the size of the cluster depends on the distance between the CH and the BS. But in such a case, there may be a scenario where the radius for a CH is large, but the energy of the node is less and also it has many nodes in its cluster. This will lead to early energy consumption of the CH, affecting the overall network performance which may be considered as a major limitation of proposed1 algorithm.

3.3.2 Proposed Solution

To resolve the problem of proposed1, we proposed another method proposed2 for cluster formation. In this work the radius of the CH depends upon all the three parameters: distance, degree and energy of the CH. The formula used is as shown in equation 3.22

$$S_i.R_{comp} = \left[1 - \left(\frac{1}{Deg\,(si) * E(i)}\left(\frac{d\,(Si,BS) - dmin}{dmax - dmin}\right)\right)\right] R^0_{comp} \tag{3.22}$$

According to this formula, firstly, the CH that is close to BS will have a cluster size large because energy consumed in data transmission to the BS is less as compared to the CHs that are away from the BS. As its energy consumed is small so it can let more nodes join its cluster and manage a cluster of big size. This helps reducing load on the CHs that are away from the BS and thus their energy consumed is reduced. Secondly, when a CH has high degree, the cluster size formed for such a CH is small, this is so because as the neighbor count for the CH is large so its need to receive energy from a large number of nodes, hence consuming a huge amount of energy. To preserve the energy of the CH, this CH will form a cluster of size small so that its energy does not drain at a rapid rate. While the CH that has small degree forms a cluster of large size to accommodate more number of nodes and manage the resource. Thirdly, the energy is considered as a very important factor while choosing the CH, similarly while forming a cluster also we consider the energy of the CH. The CH that has huge amount of energy can form a cluster with large radius while the one that has less energy would form a cluster of small size as more the number of nodes in the cluster more would be the energy consumption of the CH, leading to more energy dissipation. Thus the load on the CH should be balanced according to its parameters.

As discussed in section 3.2.2.4, various nodes overlap in more than two clusters. To remove this overlapping we use the equation 3.23

$$\frac{Dis_{CH\_BS} + Dis_{CH\_N}}{E_{CH}} \quad (3.22)$$

Comparing our proposed1 algorithm with proposed2, proposed2 performs much better in terms of the lifetime and the energy consumption.

CHAPTER 4

SIMULATION RESULTS AND ANALYSIS

Simulation is used to depict the real world system as a model. The model of the system gives the characteristics or the functionalities of the entire system. The model we have proposed is simulated and analyzed to figure out its performance. The analysis is done considering varied environmental situations. The performance of our proposed work is compared with the existing protocols in terms of the network lifetime, energy consumption and efficiency.

4.1. Simulation Setup

The simulation tool used for our work is 'MATLAB 7.7.0'. The environment for the simulation is shown in **figure 4.1**. The network is set up having 100 nodes that are located randomly over the entire field and are homogeneous. The field area has the dimensions of 150m * 150m. The BS is located at the center of the sensor field having the coordinates (75, 75). The BS has unlimited resources and skills. The data packet used in all the protocols that is transmitted between the network from the nodes to the CH and then by the CH to the BS has a size of 4000 bits. The current protocol and our proposed protocols are simulated on the same environment specifications for better analysis and results. All the parameters have been defined in the **table4.1** given below:

Table 4.1 Parameters Definition

Parameter	Value
Z * Z	150m * 150m
BS position	(75, 75)
Nr	100
b	4000 bits
E_0	1.5J
EDA	5 nJ/bit/message
E_{elec}	50 nJ/bit
ϵ_{amp}	0.0013 pJ/bit/m^4
ϵ_{fs}	10 pJ/bit/m^2

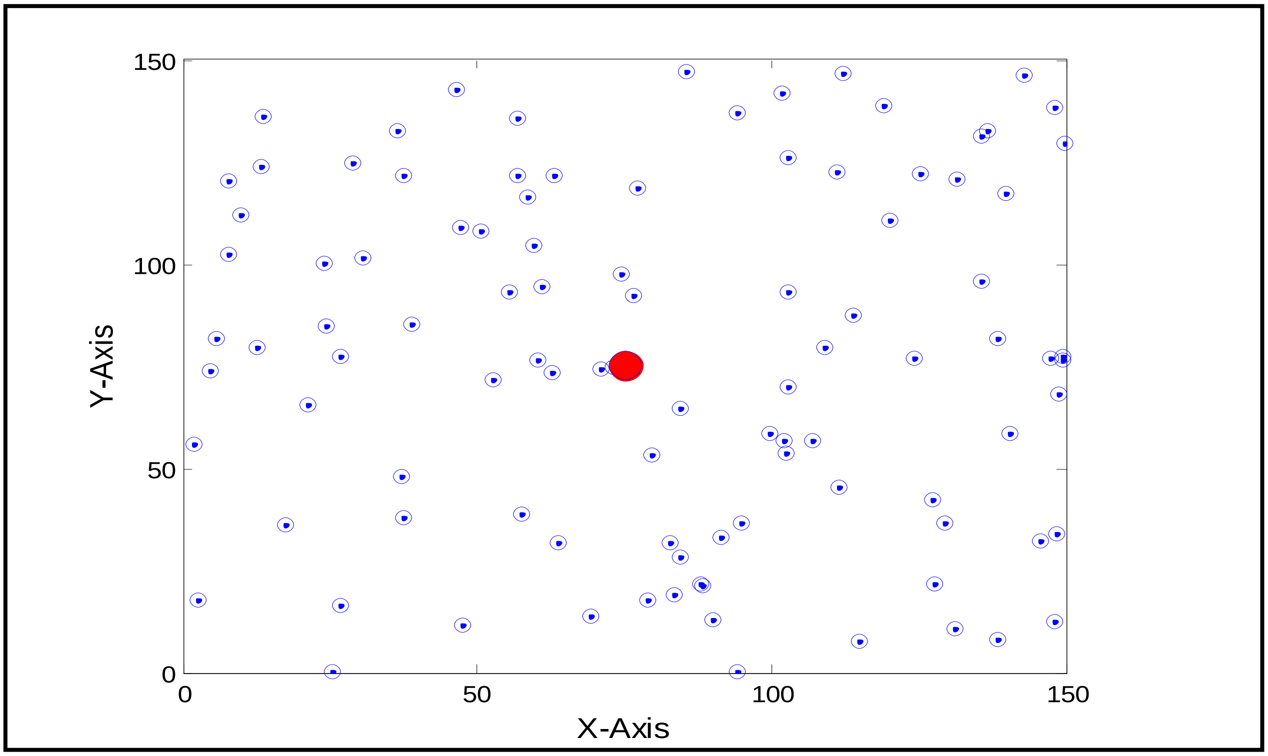

Figure4.1. Node deployment in the Network

4.2. Assessment of the Performance

This topic discusses the performance assessment of the proposed1 and proposed2 which are compared with the prevailing algorithm EECA (Energy Efficient Clustering Algorithm) [14]. The implementation is done on the MATLAB tool and all the algorithms use the same specifications of the network. The count of rounds for which the protocols are run is 1500 rounds. The following are the metrics for performance assessment:

4.2.1. Performance Metrics

a. Network Lifetime: The number of nodes that are alive in every round.
b. Remaining Energy: The total remaining energy of all the nodes in every round.
c. R1: The round number when the first node is dead.
d. R50: The round number when half nodes are dead.
e. R80: The round number when count of nodes dead is 80% of the total initial nodes.

4.2.2. Simulation Result for mobile nodes

This section presents the results achieved on simulation of proposed1 and proposed2 algorithms according to the above performance metrics. The results are then compared with the EECA-M2 [14] algorithm. The nodes here are homogeneous and have same parameters. The **figure 4.2** uptil **figures 4.4** show the results obtained on simulation of both the proposed algorithms for mobile nodes with the existing algorithm. According to the results, it can be seen that the proposed1 has balanced the network load up to a large extent but proposed2 has balanced the load even more and has optimized the lifetime.

The following figures depict the results of the simulation:

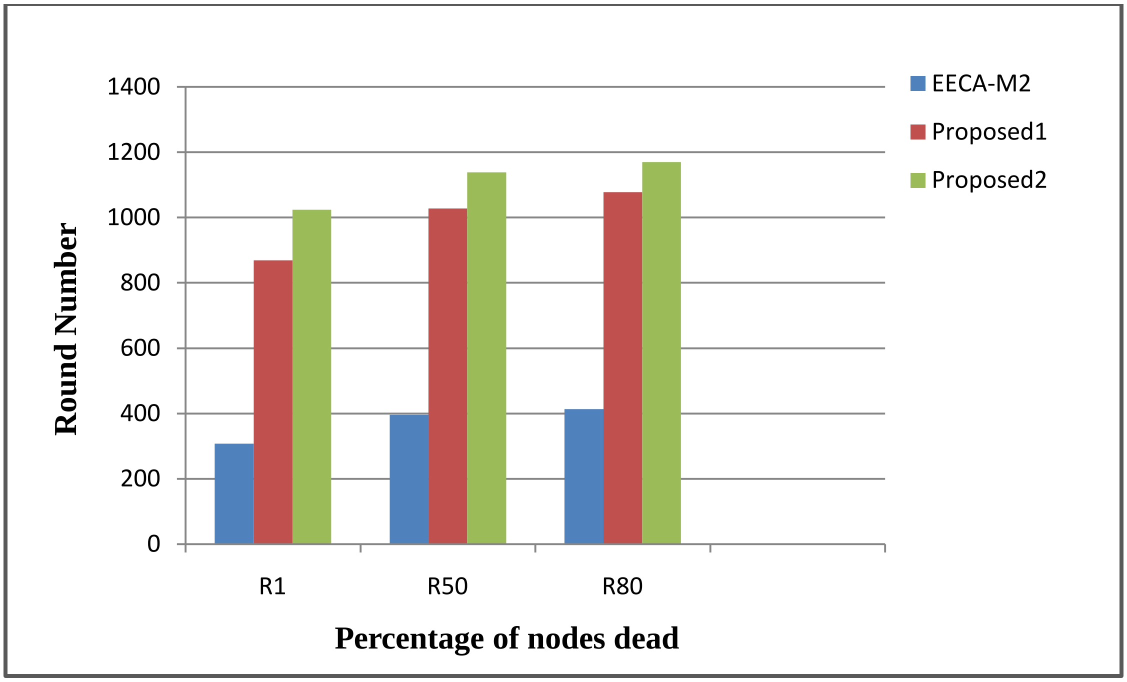

Figure4.2. Round number when nodes die

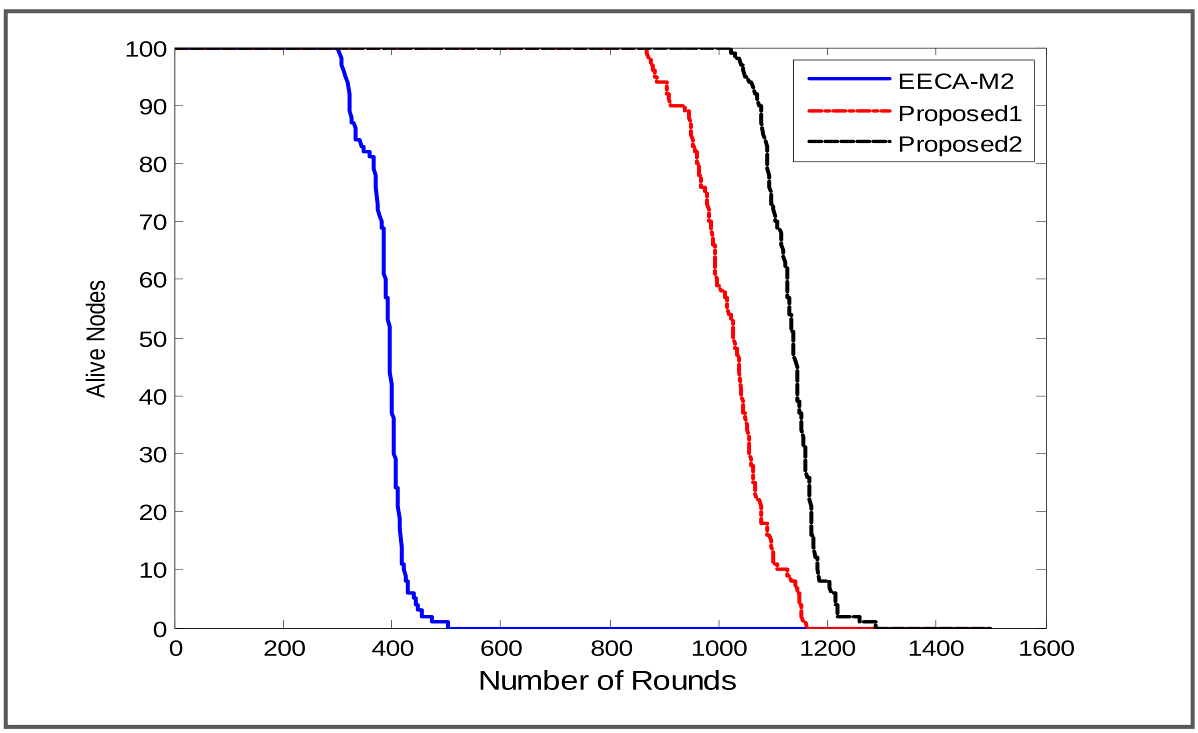

Figure4.3. Lifetime of the Network

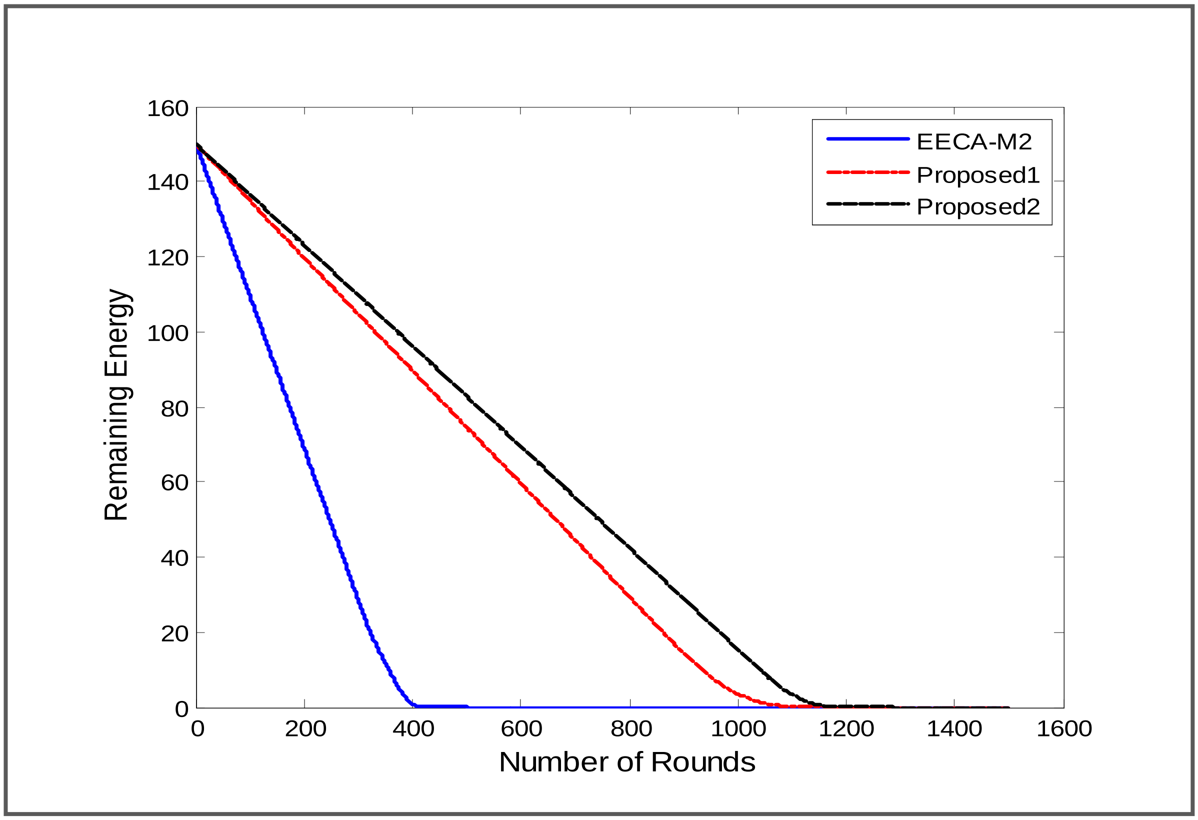

Figure4.4. Remaining energy

5.2.3. Simulation Result for fixed nodes

In this section, we discuss the results achieved on simulation of the proposed algorithm when the nodes are fixed.

The following figures depict the results of the simulation:

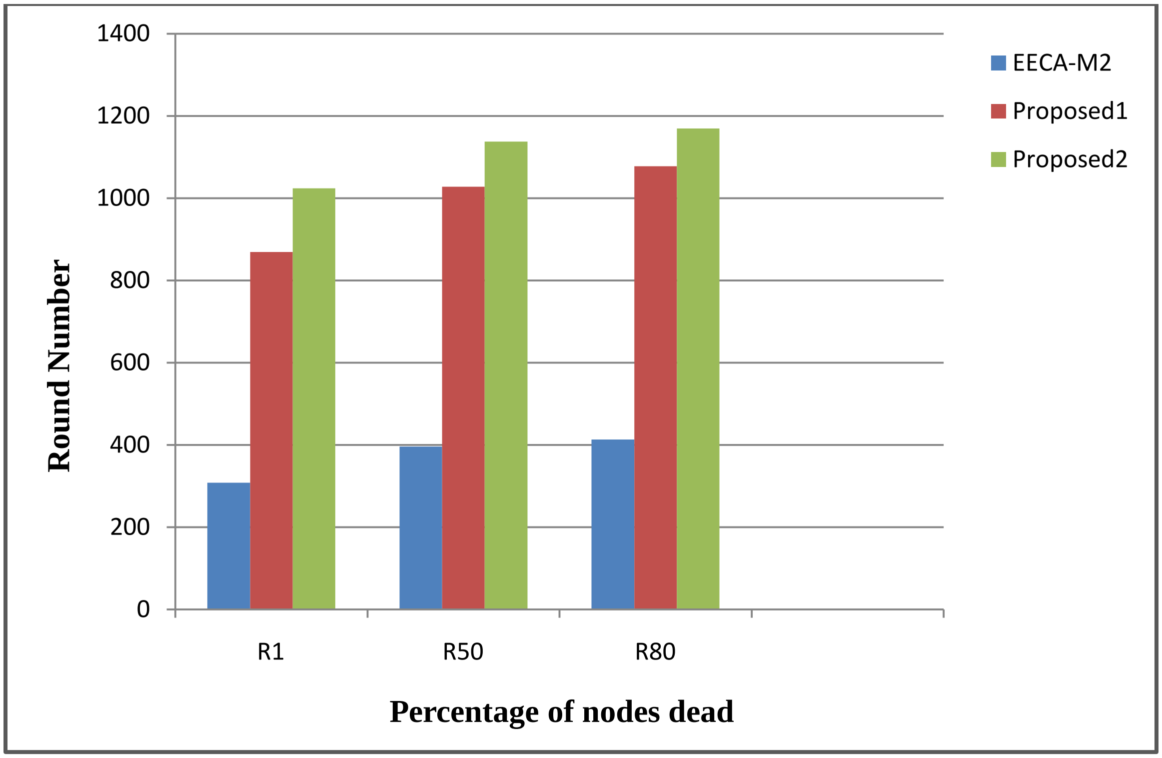

Figure4.5. Round number when nodes die

Figure 4.2 and 4.5 illustrates the round number when the first node is dead or when the half nodes are dead. Looking at the figure we can notice that the proposed2 is much stable than EECA as well as proposed1.

Figure 4.3 and 4.6 illustrates the lifetime of the network in terms of number of alive nodes in various rounds. As compared to EECA-M2, our work proposed1 has increased the number of rounds when the first node dies from 308 to 869 while in proposed2 first nodes dies at round number 1024 and when comparing EECA-F, proposed1 and proposed2 the number of rounds when the first node dies are 203, 431 and 657 respectively which gives a drastic changes in the statistics.

Figure 4.4 and 4.7 illustrates the remaining energy; this energy is the total energy of the nodes that are alive in that particular round. Thus, it shows that our protocol has balanced the energy utilization of the nodes up to an extensive amount.

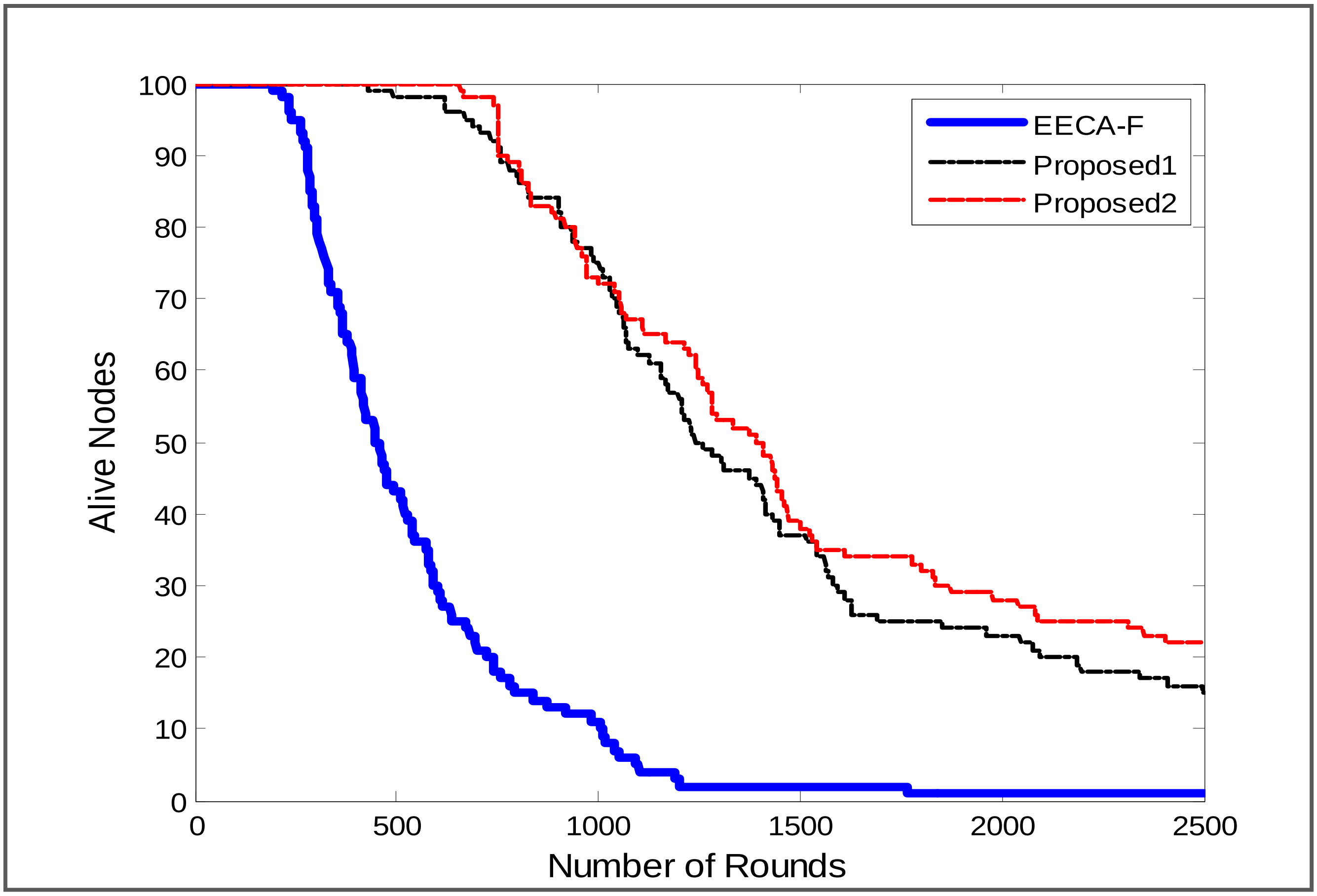

Figure4.6. Lifetime of the Network

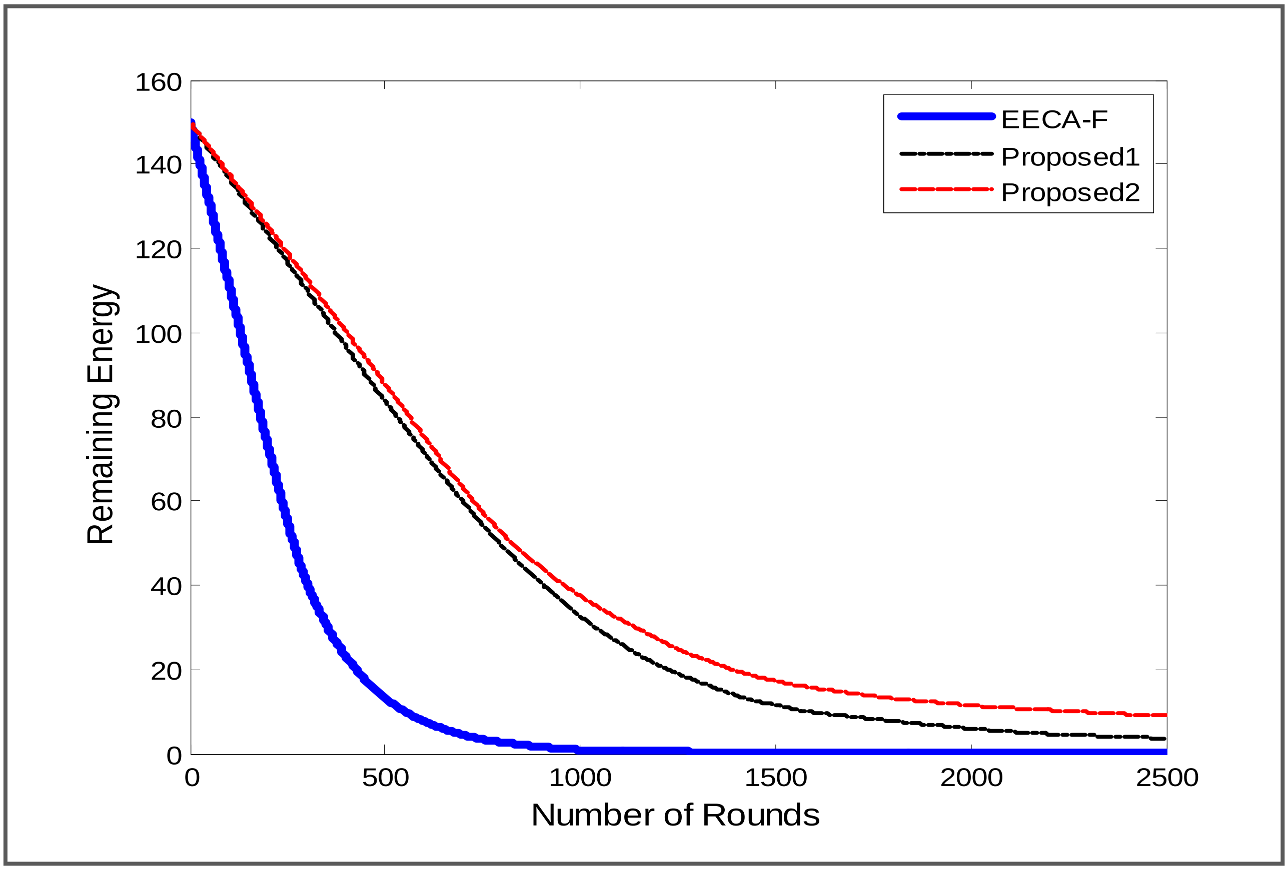

Figure4.7. Remaining energy

4.2.4. Analysis

The **table4.2** below gives a conclusion and analysis of the results achieved after simulation when the nodes are mobile.

Table4.2. Analysis for mobile nodes

Percentage of Dead Nodes	EECA-M2 (Round Number)	Proposed 1 (Round Number)	Proposed 2 (Round Number)
1%	308	869	1024
50%	396	1028	1138
80%	413	1078	1170

Here we examine that there is an enormous increase in the stability of the network, its lifetime and the remaining energy.

The **table4.3** below gives a conclusion and analysis of the results achieved after simulation when the nodes are fixed.

Table4.3. Analysis for fixed nodes

Percentage of Dead Nodes	EECA-F (Round Number)	Proposed 1 (Round Number)	Proposed 2 (Round Number)
1%	203	431	657
50%	498	1239	1390
80%	723	2096	2575

While comparing the energy remaining of all the protocols, we confer to the result that our algorithm proposed2 has a performance better than the existing algorithm EECA by about 200 to 400% when the energy of the network is compared after about 500, 1000 and 1500 rounds of the execution of the protocol.

CHAPTER 5

CONCLUSION AND FUTURE WORK

Two protocols are proposed with objective to improve the overall network lifetime and the load balanced energy consumption among the sensors of the network. This is done by taking inspiration from earlier the work carried out by researchers. Optimal CH selection still remains one of the open problems. Our work considers this problem and proposes a new solution to select cluster head using some key parameters as node degree of connectivity, distance from base station, energy and mobility. All these parameters combine to calculate the weight for each node which needs to be minimum for a node for it to become the CH for the ongoing round. Each parameter is given a priority while computing the weight. This weight helps to make a parameter more important than the other while choosing the CHs. After cluster head nodes are selected, the clusters are formed according to the cluster radius of each CH. This radius is decided depending upon the energy remaining, degree and the distance between the BS and the CH. This radius lets the load balance among the clusters in the network. Since, the size of the cluster depends upon the current capabilities of the CH, so the CH is not loaded with huge number of nodes in its cluster. This balancing of the load leads to more efficient network and hence a growth in the lifetime of the network. The results obtained by the simulation of our work are compared with the existing protocol EECA and shows that our protocol is successful in optimizing the network lifetime.

Future work can be done using the same proposed technique on the network that is heterogeneous in nature and increase its efficiency.

REFERENCES

[1] Chong, Chee-Yee, and Srikanta P. Kumar. "Sensor networks: evolution, opportunities, and challenges." Proceedings of the IEEE 91.8 (2003): 1247-1256.

[2] Xu, Ning. "A survey of sensor network applications." IEEE Communications Magazine 40.8 (2002): 102-114.

[3] Chaudhary, D. D., S. P. Nayse, and L. M. Waghmare. "Application of wireless sensor networks for greenhouse parameter control in precision agriculture." International Journal of Wireless & Mobile Networks (IJWMN) Vol 3.1 (2011): 140-149.

[4] Akyildiz, Ian F., et al. "Wireless sensor networks: a survey." Computer networks 38.4 (2002): 393-422.

[5] Sharawi, Marwa, et al. "Routing wireless sensor networks based on soft computing paradigms: survey." International Journal on Soft Computing, Artificial Intelligence and Applications (IJSCAI) 2.4 (2013): 21-36.

[6] Sharawi, Marwa, et al. "Routing wireless sensor networks based on soft computing paradigms: survey." International Journal on Soft Computing, Artificial Intelligence and Applications (IJSCAI) 2.4 (2013): 21-36.

[7] Shen, Xingfa, Zhi Wang, and Youxian Sun. "Wireless sensor networks for industrial applications." Intelligent Control and Automation, 2004. WCICA 2004. Fifth World Congress on.Vol. 4. IEEE, 2004.

[8] Ahmad Sardouk, L. M. Boulahia, R. Rahim and D. Gaiti, “Data Aggregation in WSNs: A Survey,” in Material Science Reasearch, vol. 4, no. 4.

[9] Han, Jiawei, and MichelineKamber. "Data mining: concepts and techniques (the Morgan Kaufmann Series in data management systems)." (2000).

[10] Zhao, Xinyuan, and Neng Wang. "An unequal layered clustering approach for large scale wireless sensor networks." Future Computer and Communication (ICFCC), 2010 2nd International Conference on.Vol. 1.IEEE, 2010.

[11] Li, Wei, and Guangting Chen. "Energy efficient clustering algorithm in wireless sensor networks." Wireless, Mobile and Multimedia Networks, 2006 IET International Conference on. IET, 2006.

[12] Gao, Qinghai, Junshan Zhang, and Bryan Larish. "Energy balancing in coalition-based multi-hop wireless sensor networks." Military Communications Conference, 2006. MILCOM 2006. IEEE. IEEE, 2006.

[13] Boyinbode, Olutayo, Hanh Le, and Makoto Takizawa. "A survey on clustering algorithms for wireless sensor networks." International Journal of Space-Based and Situated Computing 1.2-3 (2011): 130-136.

[14] Guiloufi, AwatefBenfradj, NejehNasri, and AbdennaceurKachouri. "Energy-efficient clustering algorithms for fixed and mobile Wireless Sensor Networks." Wireless Communications and Mobile Computing Conference (IWCMC), 2014 International. IEEE, 2014.

[15] Kumar, G. Santhosh, Vinu Paul, and K. Poulose Jacob. "Mobility metric based leach-mobile protocol." Advanced Computing and Communications, 2008. ADCOM 2008. 16th International Conference on. IEEE, 2008.

[16] Deng, Shaozhi, Jie Li, and L. Shen. "Mobility-based clustering protocol for wireless sensor networks with mobile nodes." Wireless Sensor Systems, IET 1.1 (2011): 39-47.

[17] Kwak, Byung-Jae, Nah-Oak Song, and Leonard E. Miller. "A canonical measure of mobility for mobile ad hoc networks." Military Communications Conference, 2003. MILCOM'03. 2003 IEEE. Vol. 2. IEEE, 2003.

[18] Li, Chengfa, et al. "An energy-efficient unequal clustering mechanism for wireless sensor networks." Mobile Adhoc and Sensor Systems Conference, 2005. IEEE International Conference on. IEEE, 2005.

[19] Zytoune, Ouadoudi, DrissAboutajdine, and Mehdi Tazi. "Energy balanced clustering algorithm for routing in heterogeneous wireless sensor networks."I/V Communications and Mobile Network (ISVC), 2010 5th International Symposium on. IEEE, 2010.

[20] Sen, Feng, Qi Bing, and Tang Liangrui. "An improved energy-efficient pegasis-based protocol in wireless sensor networks." Fuzzy Systems and Knowledge Discovery (FSKD), 2011 Eighth International Conference On. Vol. 4. IEEE, 2011.

[21] Heinzelman, Wendi B., Anantha P. Chandrakasan, and HariBalakrishnan. "An application-specific protocol architecture for wireless microsensor networks." Wireless Communications, IEEE Transactions on 1.4 (2002): 660-670.

[22] Heinzelman, Wendi Rabiner, AnanthaChandrakasan, and HariBalakrishnan. "Energy-efficient communication protocol for wireless microsensor networks." System sciences, 2000. Proceedings of the 33rd annual Hawaii international conference on IEEE, 2000.

[23] Beiranvand, Zahra, Ahmad Patooghy, and Mehdi Fazeli. "I-LEACH: An efficient routing algorithm to improve performance & to reduce energy consumption in Wireless Sensor Networks." Information and Knowledge Technology (IKT), 2013 5th Conference on IEEE, 2013.

[24] R. K. Yadav and A. Jain, "CHATSEP: Critical heterogeneous adaptive threshold sensitive election protocol for Wireless Sensor Networks," Advances in Computing, Communications and Informatics (ICACCI, 2014 International Conference on, New Delhi, 2014, pp. 80-86.

[25] Soua, Ridha, and Pascale Minet. "A survey on energy efficient techniques in wireless sensor networks", Wireless and Mobile Networking Conference (WMNC), 2011 4th Joint IFIP IEEE, 2011.

[26] Li, Chengfa, et al. "An energy-efficient unequal clustering mechanism for wireless sensor networks."MobileAdhoc and Sensor Systems Conference 2005, IEEE International Conference on IEEE, 2005.

[27] Malhotra, Prerna, and Ajay Dureja. "A Survey of Weight-Based Clustering Algorithms in MANET", IOSR Journal of Computer Engineering 9.6 (2013): 34-40.

[28] M. Chatterjee, S. Das and D. Turgut, “WCA: A weighted clustering algorithm for mobile ad-hoc networks,” Cluster Computing, 2002, vol. 5, no. 2, pp. 193-204.

[29] Singh, Santar Pal, and S. C. Sharma. "A survey on cluster based routing protocols in wireless sensor networks", Procedia Computer Science 45 (2015): 687-695.

[30] Sahana, Sudipta, SoumyabrataSaha, and SuparnaDasGupta. "Weight Based Hierarchical Clustering Algorithm for Mobile Ad hoc Networks", Procedia Engineering 38 (2012): 1084-1093.

[31] KiranMaraiya, Kamal Kant, Nitin Gupta, “Wireless Sensor Network: A Review on Data Aggregation”, IJSER, April- 2011.

[32] M. Chatterjee, S. Das and D. Turgut, “WCA: A weighted clustering algorithm for mobile ad-hoc networks,” Cluster Computing, 2002, vol. 5, no. 2, pp. 193-204.

[33] Aissa, Mohamed, AbdelfettahBelghith, and Khalil Drira. "New strategies and extensions in weighted clustering algorithms for mobile ad hoc networks". Procedia Computer Science 19 (2013): 297-304.

[34]Kumarawadu, Priyantha, et al. "Algorithms for node clustering in wireless sensor networks: A survey." Information and Automation for Sustainability, 2008 ICIAFS 20084th International Conference on IEEE, 2008.