Framework for Energy-Efficient Clustering Approach for Wireless Sensor Networks

Doctoral Thesis

Submitted by

Vipin Pal

Research Scholar

ID: 2010REC203

Under Supervision of

Dr. R. P. Yadav

Professor, Dept. of ECE, MNIT Jaipur

Dr. Girdhari Singh

Associate Professor, Dept. of CE, MNIT Jaipur

Department of Electronics & Communication Engineering

Malaviya National Institute of Technology Jaipur

February 2016

Framework for Energy-Efficient Clustering Approach for Wireless Sensor Networks

Submitted by

Vipin Pal

Research Scholar

ID: 2010REC203

Under Supervision of

Dr. R. P. Yadav

Professor, Dept. of ECE, MNIT Jaipur

Dr. Girdhari Singh

Associate Professor, Dept. of CE, MNIT Jaipur

Submitted in fulfillment of the requirements of the degree of Doctor of Philosophy

to the

Department of Electronics & Communication Engineering

Malaviya National Institute of Technology Jaipur

February 2016

Certificate

This is to certify that the thesis entitled "**Framework for Energy-Efficient Clustering Approach for Wireless Sensor Networks**" being submitted by **Vipin Pal** is a bonafide research work carried out under our supervision and guidance in fulfillment of the requirement for the award of the degree of **Doctor of Philosophy** in the Department of Electronics & Communication Engineering, Malaviya National Institute of Technology Jaipur. The matter embodied in this thesis is original and has not been submitted to any other University or Institute for the award of any other degree.

(Dr. R P Yadav)
Professor

(Dr. Girdhari Singh)
Associate Professor

Declaration of Authorship

I, Vipin Pal , declare that this thesis titled, 'Framework for Energy-Efficient Clustering Approach for Wireless Sensor Networks' and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

“You cannot believe in God until you believe in yourself.”

- Swami Vivekananda

Acknowledgements

Information and communication are two of the most important strategic issues for the development of a nation. In this modern era of socio-technical growth, computer networking has a substantial role for a nations prosperity. Being a PhD scholar in Electronics and Communication Engineering Department of Malaviya National Institute of Technology, Jaipur it was my privilege to be familiar with borderless communication and information environment of computer networks. I feel highly indebted to my supervisors Professor R P Yadav and Dr. Girdhari Singh for considering me in their research group and providing me content to explore the research area. I am deeply thankful to them for their constant faith in my ability and encouragement towards achieving my goal.

Besides my supervisors, I express my sincere gratitude toward Dr. Vineet Sahula (Prof. and Head of ECE Dept.) and Dr. Md. Salim for their steady support. I would like to show appreciation to my PhD committee members: Dr. Vijay Janyani, Dr. K K Sharma, Dr. K R Niyazi, Dr. Ghanshayam Singh and Dr. Ritu Sharma for their valuable suggestions which always guided me to provide right direction. I would like to mention the pioneer character Dr. V Sinha (Distinguished and Emeritus Professor) Malaviya National Institute of Technology, Jaipur for sharing their experience from his vast acquaintance. I appreciate all faculty members of ECE and Computer Engineering Department of Malaviya National Institute of Technology, Jaipur.

I am also grateful to Prof. I. K. Bhatt, Director Malaviya National Institute of Technology, Jaipur for providing me the necessary infrastructure and facilities which were necessary to carry out this work.

I am enormously endeavored to my family for their constant support which gave me mental strength and much needed will power to complete my work. Last but not the least; I give my warm thanks to my colleagues and friends for their help, moral support, friendly attitude and valuable suggestions. Without them all that I have achieved would not have been possible.

I am thankful to Ministry of Human Resource and Development (MHRD), India for financial support throughout my PhD period.

Abstract

The work of presented thesis contributes to develop energy efficient protocols for wireless sensor networks. In wireless sensor networks, nodes operate in the application area with severe resource constraints like limited battery power, less memory, low bandwidth and computing ability. Effective and energy efficient protocols have been the primary design issue for wireless sensor networks. Clustering algorithms are considered as energy efficient approach and help to provide reliability and scalability.

For data centric wireless sensor networks, a load balanced clustering algorithm is proposed that has network adaptive round-time and divides the network area into two parts, border area and inner area, for cluster head selection. Round-time is calculated according to number of nodes alive that solves fixed round-time problem. In proposed solution only inner area nodes can contend for cluster head role. Proposed solution has more chances of a selected cluster head to be near to the center of cluster. Simulation results demonstrate that proposed solution has lower node death and energy consumption rate. More data units are received successfully at base station in case of proposed solution compared to LEACH.

Heterogeneity of nodes improves performance of clustering algorithms. We exploit heterogeneity of nodes either caused by network dynamics in an initial homogeneous network or initially few nodes have more energy than other. Proposed solution for cluster head selection considers both residual energy of node and average energy of network. Nodes with higher energy have high probability to be selected as cluster head in early rounds of an epoch and also have more chances of selection as cluster head in an epoch to load balance the network. Simulation results suggest that proposed solution has better network lifetime and data gathering at base station compared to LEACH and SEP.

We have also analyzed the effect of cluster size on the performance of clustering algorithms. Small-sized clusters consume more energy than large-sized clusters because node of small-sized clusters have more time slots in a data transmission phase than node of large-sized clusters. A clustering approach is proposed that have balance-sized clusters depending upon a distance threshold. Cluster formation is accomplished in two phases. At the end of first phase, few nodes are left un-clustered. These un-clustered nodes join the best possible cluster to have better cluster quality. Simulation results show

that proposed solution of balance-sized clusters has better cluster quality and prolonged network lifetime compared to LEACH.

Wireless sensor networks should produce reliable information about the application area. Reliability state of a wireless sensor network can be measured as the coverage area of deployed network and data redundancy in the data due to overlapped sensed area of nodes. Dynamic nature of wireless sensor networks make it imperative to find transition of reliable state to unreliable state. We use genetic algorithm based approach to find the minimum number of nodes (random) which sense almost complete area with desired minimum redundancy in the data. Results are useful for both clustering approaches and sleep scheduling algorithms.

Contents

List of Figures

List of Tables

Abbreviations

Acronym	**W**hat (it) **S**tands **F**or
SOSUS	**So**und **Su**rveillance **S**ystem
NOAA	**N**ational **O**ceanographic and **A**tmosphere **A**dministration
DARPA	**D**efense **A**dvanced **R**esearch **P**roject **A**gency
DSN	**D**istributed **S**ensor **N**etworks
MEMS	**M**icro-**e**lectro-**m**echanical-**s**ystems
WLAN	**W**ireless **L**ocal **A**rea **N**etwork
TDMA	**T**ime **D**ivison **M**ultiple **A**ccess
MAC	**M**edium **A**ccess **C**ontrol
LEACH	**L**ow **E**nergy **A**daptive **C**luster **H**ierarchy
CSMA	**C**arrier **S**ense **M**ultiple **A**ccess
S-MAC	**S**ensor **M**edium **A**ccess **C**ontrol
RTS	**R**equest-**t**o-**S**end
CTS	**C**lear-**t**o-**S**end
FRTS	**F**uture **R**equest-**t**o-**S**end
LMAC	**L**ightweight **M**edium **A**ccess **C**ontrol
CSMA/CA	**C**arrier **S**ense **M**ultiple **A**ccess/**C**ollision **A**voidance
TRAMA	**Tra**ffic-adaptive **M**edium **A**cess **P**rotocol
FlexiTP	**Flexi**ble-Schedule-based **T**DMA **P**rotocol
DRAND	**D**istributed **Rand**omized TDMA Scheduling
STEM	**S**parse **T**opology and **E**nergy **M**anagement
DSSP	**D**yanimc **S**leep **S**cheduling **P**rotocol
TTS	**T**wo **T**ier **S**cheduling
RSS	**R**eceived **S**ignal **S**trength
SEP	**S**table **E**lection **P**rotocol

DEEC	**D**istibuted **E**nergy-**E**fficient **C**lustering
EEHC	**E**nergy **E**fficient **H**eterogeneous **C**lustering
ERP	**E**nergy-aware **R**outing **P**rotocol
EBUC	**E**nergy-**B**alanced **U**nequal **C**lustering
PSO	**P**article **S**wam **O**ptimization
FCM	**F**uzzy **C**-**M**eans
ACO	**A**nt **C**olony **O**ptimization
IWD	**I**ntelligent **W**ater **D**rop
GA	**G**netic **A**lgorithm
ADRP	**A**daptive **D**ecenteralized **R**e-clustering **P**rotocol
HCR	**H**ierarchical **C**luster-based **R**outing
PACT	**P**ower **A**ware **C**lustered **T**DMA
EEPSC	**E**nergy-**E**fficient **P**rotocol with **S**tatic **C**lustering
EDASC	**E**nergy-efficient **D**ata **A**ggregation Protocol based on **S**tatic **C**lustering
QAC	**Q**oS-based **A**daptive **C**lustering
NS-2	**N**etwork **S**imulator (version-2)
OMNET	**OMNET**
OPNET++	**OPNET**
GloMoSim	**Glo**bal **Mo**bile Information System **Sim**ulator
OTcl	**O**bject-oriented **T**ool **C**ommand **L**angauge
MATLAB	**MAT**rix **LAB**oratory
GPS	**G**lobal **P**ositioning **S**ystem
NP-HARD	**N**on-deterministic **P**olynomial-time **hard**
TSP	**T**ravelling **S**alesman **P**roblem

Dedicated to the memories of my grandmother and my cousin and to my family for their love, endless support and encouragement…

Chapter 1

Introduction

Recent advancements in the field of semiconductor, networking and other related technologies are making rapid growth of wireless sensor networks. Advanced wireless sensor networks have lower maintenance and deployment cost, better battery power and are less prone to failure. Therefore, they are finding application in military, home, industry, health, environment and many more areas.

1.1 Introduction

As like many advanced technologies, the requirements of military operations gave birth to wireless sensor networks in the 1950s during the Cold War [1, 2, 3, 4]. The first wireless sensor network was Sound Surveillance System (SOSUS) developed by the United States Military. The system used acoustic sensors (hydrophones) which were deployed strategically in the Atlantic and Pacific oceans to detect and track Soviet submarines. The system is currently used by the National Oceanographic and Atmosphere Administration (NOAA) to monitor ocean events.

Modern wireless sensor networks started around 1980 at the Defense Advanced Research Project Agency (DARPA) with the Distributed Sensor Networks (DSN) program [4]. The combination of high cost and bulky sensor nodes restrict the application area of wireless sensor networks to military, heavy industry and few academies. Advancement in micro-electro-mechanical-systems (MEMS) provides inexpensive low-power sensor nodes with wireless technologies that enhance the stationing of wireless sensor networks into a broader range of applications such as surveillance, smart home, healthcare, agriculture, automation, industry, traffic management, habitat monitoring, disaster detection and management, environment monitoring and various others.

A wireless sensor network comprises of numerous densely placed sensor nodes in the nucleus of application area or very near it [5, 6]. Sensor node is the primary element of wireless sensor network which is constructed of computation, sensing and wireless communication components with an on-board battery. Hence, in the application area, the deployed nodes have the ability of processing, sensing and wireless communication. Since a single node does not have adequate skills of completing the task therefore, they work together to complete the task. The application area is sensed by the sensor nodes and that sensed data is transmitted to the base station directly or through multi-hop communication. Figure 1.1 demonstrates framework of a wireless sensor network. Base station accumulates all the data and end users access the heaped data directly or via a remote connection. Base station plays as a gateway between end user and deployed sensor nodes.

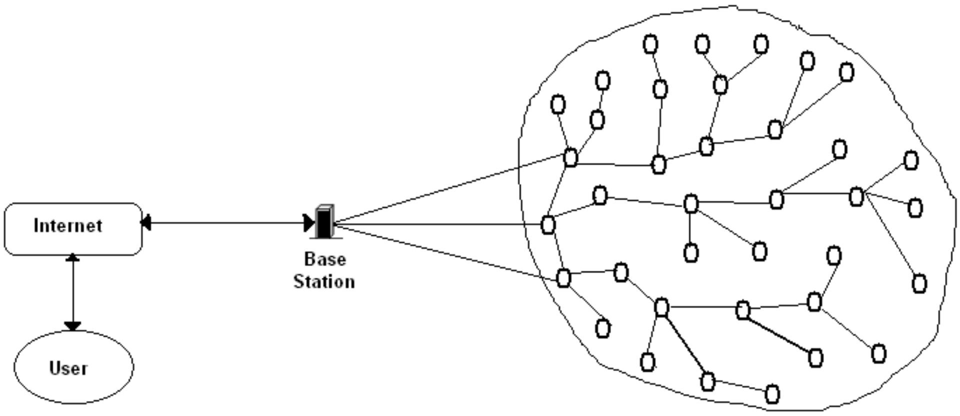


FIGURE 1.1: Wireless Sensor Networks.

Wireless sensor networks operate in the application area without any human intervention once they are deployed. Nodes coordinate with each other (sometime to some centralized unit) to start or maintain the network. So wireless sensor networks are infrastructure less network, hence are different from cellular networks and WLANs which are infrastructure oriented networks. Even though, wireless sensor networks and ad-hoc networks are infrastructure less networks, they are different in many aspects because of some unique features of wireless sensor networks [5, 7, 8, 9, 10] as follows:

1. Number of nodes in wireless sensor networks is much higher compared to devices of ad-hoc networks; and number of nodes in a wireless sensor network can fluctuate from few nodes to hundreds of nodes. Therefore wireless sensor networks are more scalable to ad-hoc networks and require a more scalable solution.

2. Energy is scarce in both networks. Recharging or replacement of batteries of sensor nodes is not quite possible; while it is possible for ad-hoc networks. Devices of ad-hoc networks carry more energy compared to sensor nodes. Energy consideration is much higher in wireless sensor networks. Thus energy efficiency is the foremost design issue for algorithms of wireless sensor networks.

3. In most of the application, sensor nodes have no human intervention after the deployment in the environment. Therefore, nodes should configure and maintain the network autonomously. While, there are continuous human interventions in ad-hoc networks.

4. Sensor nodes are always exposed to harsh environment. Therefore sensor nodes are more prone to failure. Hence in wireless sensor networks, probability of topology change is much higher. Therefore, sensor networks should be more fault-tolerant and capable of adapting to the changes than ad hoc networks.

5. Sensor nodes are deployed more densely as compared to density of devices in ad-hoc networks. Depending on the application, density of sensor network changes accordingly, for example, military applications require data of high accuracy so have high density.

1.2 Motivation

Wireless sensor networks have extensive range of potential applications. Applications of wireless sensor networks can be grouped mainly into two classes according to their operational pattern: data gathering and event driven [11]. Data gathering applications require periodically information of application area at the base station. Nodes continuously sense and send the information to base station. In these applications, data transfer rate is almost constant throughout the working of network. Examples of data gathering application are health monitoring [12], humidity and temperature monitoring [13], habitat monitoring [14]. In an event-driven application when some event is triggered in the application area, nodes send the data to the base station. In these applications, network has a rugged data transfer when an event is detected. A chemical spill monitoring [15], automobiles [16], industry [17], intrusion detection [18], critical information infrastructure protection [19] are the examples of event-driven applications.

Sensor nodes work in harsh\remote field without much human intervention. These sensor nodes have resource limitations like finite battery power, tight computational power and less memory capacity. Limited battery power of nodes make energy of nodes as the most valuable resource because the overall operation and performance of network depends upon the life of sensor nodes. It is utterly impossible to recharge or replace battery of sensor nodes; therefore efficient energy utilization of sensor nodes is the central design issue for researchers. So there is more prominence on efficient energy utilization of sensor nodes, from circuitry of sensor nodes to application level to protocols of networks, while keeping impeccable performance of network. Deployed network supposed to be covered the entire area so large population of sensors are expected. Consequently, applied energy efficient protocols should offer high reliability in order in order to have longer network lifetime.

Clustering algorithms have been applied to diminish energy consumption of sensor nodes and also satisfy the scalability objective with prolonged network lifetime in large scale Wireless sensor network environment [20, 21, 22]. Clustering is an idea to group nodes in independent sets and to select a cluster head node for each cluster. Clustering algorithms perform data aggregation/fusion to decrease the number of transmitted message to the base station and lower the energy consumption within a cluster. Applications that require data aggregation (e.g. computing the maximum temperature around a large area) are natural subscriber of clustering. Single-tier architecture is not scalable for a larger set of sensors covering a wider area because the sensor nodes are not capable of so long communication. In multi-hop communication the relay nodes are heavily loaded. Clustering algorithms are particularly useful for applications that require scalability to hundreds or thousands of nodes.

Performance of clustering algorithms heavily leans on the unique characteristics like, cluster head selection, structure of clusters, re-clustering, time span for re-clustering (round time), and more. Cluster head selection plays a vital role for clustering algorithms because a cluster head node controls complete functioning of cluster. An energy deficient cluster head node cannot carry the responsibility of cluster and degrades performance. A cluster without a cluster head node will not be in the position to send its information to base station and that will cause a sensing hole in the field. Therefore, a cluster head failure will degrade performance of network. Number of nodes in each cluster and position of cluster head in cluster defines the performance of clustering algorithm. There is requirement of re-clustering, i.e. to construct cluster and to select cluster head, to load balance the network because a cluster head node works much higher than other nodes and are more liable to consume their energy early. Re-clustering selects a new set of cluster head nodes. Time span for re-clustering is an important issue because a long round time will drain a high amount of energy from cluster head nodes while a short round time will consume more energy in clustering overhead.

In this thesis, lifetime of network is considered as the prime performance parameter. An increased lifetime of network will provide extra time to sense the field and to send the data to base station and hence there will be an increase in data packets received at base station. The work of this thesis puts an emphasis on round time for re-clustering, selection of cluster head nodes and structure of clusters.

1.3 Clustering in Wireless Sensor Networks

Structure of a wireless sensor networks in the application area can be considered as flat or hierarchical. In a flat network, all nodes are considered identical and all nodes participate in decision-making for network protocols, like routing. In a hierarchical network, few nodes are considered special. The selected special nodes are considered the backbone of the network [7]. Most of the decisions in the field are taken by these special nodes. It is easy to implement and maintain a flat network as long as it is small. But as the network augments, a flat network is difficult to manage. Hierarchical networks take advantage of hierarchy to maintain the large network.

a. Clustering

Dividing the network into independent clusters and choosing one cluster head for each cluster is an idea of hierarchical network [8, 23, 24]. Clusters are subsets of nodes, and the union of these subsets provides the complete graph of the whole network. If a graph

G= (V, E) represents a network, there is a set of subsets of nodes V_i,i=1,2,...,n such that $U_{i=1,2,\ldots,n}$ V_i=V. Partitioning of network (V) into several clusters does not provide hierarchy in network, i.e., all nodes are equal. For each subset of nodes (cluster) a special node c_i (cluster head) is selected. Cluster head represents the subset and is responsible for various tasks. The concept of clusters and cluster head provides the hierarchy in network and is called clustering.

In clustering, division of network into cluster and selection of cluster head is done either by the sensor nodes or by a centralized point of network. Selected cluster head may be just a sensor node or a resource rich sensor node. Nodes in clusters may be fixed or variable. Nodes get the impression of the environment and send the data to respective cluster head. Cluster head node is the point of aggregation of data flow in the network. Aggregation [25] is the process of computing a smaller representation of a number of messages that is equivalent to the contents of all the individual messages. Aggregated data is forwarded in the network to reach at base station.

b. Classification

Clustering algorithms can be categorized into *centralized* and *distributed* ones. In distributed clustering algorithms, nodes communicate locally to gather information from other nodes to select cluster head and to structure the clusters. In centralized algorithms, all nodes transmit their information to base station. Base station provides information of clusters and cluster heads to the network by using the network information. Scalability of distributed algorithms is higher than that of the centralized algorithms because determining the global view of network limits the scalability of centralized algorithms. Clustering algorithms can be classified in accord with cluster formation as *static* and *dynamic*. In static clustering algorithms, clusters are organized once at the inception of the network. Clusters are re-formed periodically in dynamic clustering algorithms. Static algorithms have advantage of avoiding overhead of frequent re-clustering but are not scalable. According to the characteristics of sensor nodes clustering algorithms can be categorized as *homogenous clustering* and *heterogeneous clustering*. Heterogeneous clustering algorithms have hierarchy of nodes according to their available resources. Resource rich nodes have extra importance in the network. Homogenous clustering schemes do not distinguish nodes even they are different, i.e., all nodes are treated equal.

c. Design issues of Clustering Algorithms

Clustering algorithms are effective in achieving the desired goals of wireless sensor networks. There are several key design issues for clustering algorithms to take care of.

- Load balancing is an important issue for clustering algorithm. Nodes have different duties in clusters and hence nodes consume different amount of energy. Cluster head receives data from all nodes in the clusters, aggregate the collected data and send it to base station directly or via other cluster heads as relay. Thus cluster head nodes are more prone to consuming their energy early; hence rotation of the role of cluster head among nodes is important for load balancing. Allotement of equal number of nodes in each cluster is desired for load balancing. Load balancing causes longevity of network.
- Sensor nodes are always exposed to harsh environment and thus nodes are more prone to failure. Malfunctioning of few nodes should not affect the overall performance of clustering algorithm. Thus clustering algorithms should be fault-tolerance. Re-clustering is the most spontaneous way to overcome node failure. A back-up cluster head is an option to overcome cluster head failure. Rotation of the role of cluster head among nodes also provides fault-tolerance. Fault-tolerance increases the reliability of network.
- Synchronization of nodes is required for maintaining the effectiveness of clustering algorithms. Clustering algorithms mostly use slotted transmission scheme (like TDMA) that allows nodes to transit to sleep state to save energy. Hence synchronization is required to start and maintain transmission.
- Depending on the application, the number of clusters is an important issue for overall performance of networks.

d. Energy Saving by Clustering

To increase the lifetime of network, clustering algorithms should reduce energy consumption caused by collision, over hearing, idle listening and control packets. Collision happens when two or more nodes simultaneously send data to a common node. It results in retransmission of packet. Overhearing is caused when a node receives a packet that is not intended for it because receiving process also consumes energy. In idle listening, a node is ready to receive but none sends. Transmission of control packets to maintain the network also consumes energy.

In clustering algorithms, time slotted transmission is preferred. Nodes send/receive data according to allotted time slot that avoids collision. When nodes do not send or receive, they transit to sleep state; hence reduce idle listening and overhearing. Clustering algorithms reduce control packets by using some pre-defined conditions. Inter-cluster collision is avoided by using different codes.

Communication between two nodes consumes more energy than other operations performed by nodes. Both transmitting and receiving of message consumes energy. Amount of energy consumed rely on the distance between two nodes. Clustering divides the network into clusters and hence nodes avoid long distance communication to base station; as now they send data to cluster head. Therefore, energy consumption by communication is reduced by clustering algorithms.

e. Challenges

There are several key factors that determine the effectiveness of a clustering algorithm.

- After how much time re-clustering should be performed? A long round time will drain a great amount of energy from cluster head nodes while a short round time will burden the network by clustering overhead.
- Number of cluster head nodes is another challenge for clustering algorithms. A high number of cluster head nodes will not take advantage of clustering while a small number of cluster head nodes will increase the work of cluster head node.
- Selection of node as cluster head the affects performance of network. A node with very little remaining energy will not be able to complete the round and it will degrade the performance of network.
- Position of selected cluster head node is another challenge for clustering algorithm because intra-cluster and cluster head to base station communication depends on the position of cluster head node.
- Number of nodes in each cluster is another challenge. Equal-sized clusters have better load balanced network than uneven-sized clusters. So balanced-sized clusters are required.

1.4 Objective

According to challenges imposed by clustering algorithms and their design issues, the main objective of the present research work are as:

- To investigate effectiveness of round-time for clustering algorithm and purpose an network adaptive round-time according to network dynamics.
- To design an effective cluster formation scheme that results in reduced intra-cluster communication distance of network to load balance the network.
- To explore and exploit network heterogeneity for selection of cluster head nodes for the longevity of network lifetime.
- To investigate the effect of size of clusters and to design an clustering algorithm to have balanced size clusters to increase lifetime of network.
- To investigate and find reliability of network.

The main objective of the present research work is to prolong the network lifetime and data gathered at base station with load balancing of network.

1.5 Proposed Solution to Clustering

In this section, a brief overview of proposed solutions to challenges of clustering algorithms to achieve the desired objective is outlined. A dense uniform distribution of sensor nodes is considered. Proposed solutions are distributed in nature.

a. Load Balanced Clustering Algorithm for Data-centric Wireless Sensor Networks

A load balanced clustering algorithm is proposed that has adaptive round-time and for cluster head selection divides the network area in two parts: *Border Area and Inner Area*. Proposed solution exploits fixed round time problem of traditional clustering algorithms. In proposed solution, round time is calculated according to the number of nodes alive in the network. Initially, all nodes are alive in the network, so round time is same for early rounds. As nodes start dying, round time is changed (reduced) that rely on the number of nodes alive. A fixed round time has more frames per round for a less number of nodes in the progressed network. In proposed solution, round time is made dynamic till 50% node death because nodes have less remaining energy and a short round time will consume more energy of nodes due to frequent clustering overhead.

Intra-cluster communication distance is much higher than the distance of cluster heads to base station because communication between member nodes and cluster head is higher than the communication between cluster heads and base station. Intra-cluster communication distance depends on the location of cluster head in the cluster. A cluster with cluster head situated near to the center of cluster has low intra-cluster communication distance. The proposed solution to cluster head selection reduces intra-cluster communication distance.

Proposed solution divides the area into two sections: *border area* and *inner area*. Cluster head selection process is restricted to only inner area nodes, i.e. only inner area nodes are eligible for cluster head process. In the first phase, a network partitioning distance is approximated to partition the network. Network partitioning distance has the main impact on the performance of proposed solution because a less value will not change so much while a high value will degrade the performance. Nodes have knowledge of their location in the field by means of location detection algorithms to check their status, either they belong to border area or inner area.

In proposed solution, there are more chances of a selected cluster head to be near to the center of the cluster and hence reduce intra-cluster communication distance. Proposed solution also does load balancing. Border area nodes are always at the border of the cluster and hence consume more energy but they are not in the role of the cluster head that consumes high energy. Inner nodes are always near to cluster head hence have shorter communication distance but they are in the role of the cluster head frequently.

b. Cluster head selection by exploiting heterogeneity of sensor nodes

Clustering approach that consider heterogeneity of sensor nodes which is either due to dynamic nature of network or a percentage of the total nodes have extra energy is presented. An initial homogenous wireless sensor network does not maintain its homogeneous nature as the network starts processing because nodes expend different amount of energy due to their divergent roles in the network and different communication distance. Cluster head nodes squander much higher energy comparative to member nodes of clusters. In a cluster, nodes have different distance to cluster head so they consume different amount of energy.

Initial heterogeneous wireless sensor networks refines performance of network without requiring much increase in cost [26].Proposed solution takes advantage of network heterogeneity caused by either network dynamics or initial heterogeneity of nodes for cluster head selection. So it adapts heterogeneous environment of wireless sensor networks. Nodes with high energy have high probability to be selected as cluster head in early rounds and have more chances of selection as cluster head to balance the load in the network. Cluster head selection process considers both residual energy of node and average energy of the network.

c. Balance-Sized Cluster Solution to Extend Lifetime of Wireless Sensor Networks

Size of clusters is the main challenge to clustering algorithms. Frame length of a cluster and number of frames per round depends on the number of cluster members. Uneven cluster size makes uneven frame length and frames per round for each cluster. Small-sized clusters have short frame length and hence have more frames per round. Therefore, small-sized clusters consume more energy than the large-sized ones. [5] suggests that for clustering algorithms equal-sized clusters are desired to load balance the network.

Proposed solution for cluster formation tends to construct balanced-sized clusters. Cluster formation is done in two phases. In first phase, cluster heads have restriction on the number of member nodes that rely on nodes alive and number of cluster heads. After the completion of the first phase, few nodes are un-clustered because the nearest cluster head have already member nodes to equal the defined threshold. For the second phase, a distance threshold is considered to join the cluster head. An un-clustered node will join a cluster head that is positioned closer than the distance threshold and the number of nodes is less than cluster threshold. If two or more cluster heads satisfy the joined condition, node will join the nearest cluster head that satisfies joined condition. If none of the cluster heads satisfies the joined condition, the node will join the nearest cluster head irrespective of that the node is not satisfying the condition. Size of clusters depends on distance threshold defined for second phase of cluster formation.

d. Reliability of The Deployed Wireless Sensor Network

Deployed wireless sensor network should produce enough reliable information of the field that can reflect almost accurate properties of the phenomenon. As the network progresses, nodes die randomly in the field. So the progresses should be enough reliable for the desired application. So it is necessary to find when the deployed network is in an unreliable state. Genetic algorithm is applied to find the solution.

1.6 Thesis Organization

The rest of the thesis is compiled as follows: In Chapter 2, we review the performance of various MAC and sleep/wake-up scheduling protocols and clustering algorithms for wireless sensor networks.

Chapter 3 describes the wireless sensor network model and related assumptions made along with details of radio propagation model.

In Chapter 4, we presents a load balanced clustering algorithm for data-centric wireless sensor networks that have network adaptive round time and divides the network area for cluster head selection. We first explain our approach to have network adaptive round-time and then determine the role of nodes in clustering process according to their position in the field. We also compare our scheme with existing clustering algorithm.

In Chapter 5, we present a cluster head selection approach by exploiting the heterogeneity of nodes in the field. First we describe the process of cluster head selection of our approach and then use simulation experiments to examine the performance of our approach to compare it with existing approaches.

In chapter 6, we describe a cluster formation procedure for equal size clusters. Proposed work defines two thresholds for cluster formation. Significance of these thresholds on cluster formation is explained. Proposed solution is simulated to analyze the performance.

In chapter 7, reliability state of wireless sensor network is examined. Genetic algorithm is applied to find state of wireless sensor network.

Chapter 2

Review of Literature

Efficient energy consumption of nodes has been recommended for wireless sensor networks from manufacturing of nodes (circuitry of nodes) to protocols and algorithms used at different layers of network architecture. A node performs various operations in the field, like sensing, processing data, relaying data, communicating data and many more, which are energy consuming. Communication between two nodes is considered the most energy consuming process. Most of the proposed solutions in literature save energy of nodes by reducing energy consumption in communication.

2.1 MAC Protocols for Wireless Sensor Networks

Typical MAC protocols of wireless networks are not considered proficient for wireless sensor networks [27]. *Idle listening, overhearing, overhead* and *collision* cause unnecessary energy consumption of nodes [28, 29, 30] during the MAC protocol. In wireless medium, collision occurs when more than one node, which are not in the range of each other, send data packet to a common node which is in the range of these nodes. Collision results in the retransmission of the same packet. Wireless medium is a broadcast channel. All neighboring nodes of a sender receive packet that is intended for only one node or few nodes. A node is said to be in idle listening state if it is ready to receive packet but no one is ready to send packet. Control packets also cause energy consumption. Sensor networks have high number of nodes so control packets are also high in numbers. MAC protocols for wireless sensor networks should avoid or lower these unnecessary causes of energy consumption while providing fair performance.

WiseMAC [31, 32] protocol is based on preamble sampling technique [33] to function on WiseNET. Protocol architecture of WiseMAC is based on non-persistent CSMA. Preamble sampling technique is used to reduce the effect on power which one caused by idle listening.

The same constant period, Tw, is used by all nodes to sample the network. If a node finds the medium busy, it will continuously listen to the medium until a data packet is received or the medium becomes idle. Transmitting node adds a wake-up preamble of an equal size to sampling period before data packet that will ensure the receiver is already awaked when data portion of the packet arrives. Main problem for the protocol is long wake-up preamble that limits throughput and also causes large energy consumption. But to overcome that limitation, acknowledgement also carries information of the sampling time of other party. So, a node has updated the table with sampling time offsets of all its probable destinations. With the help of this updated information, wake-up preamble size is minimized and node sends the data at the right time. Figure 2.1 shows WiseMAC.

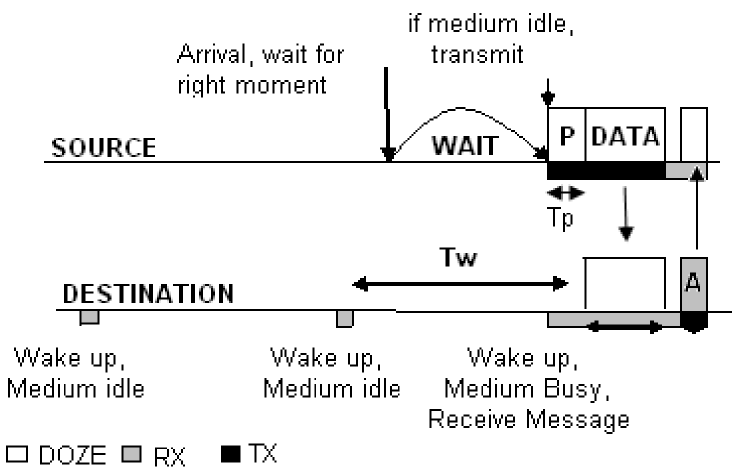


Figure 2.1: WiseMAC

Long wake-up preamble is always the first step of communication between two nodes. As the communication proceeds, exchange of sampling time information reduces size of wake-up preamble. For a higher traffic, the length of the wake-up preamble will be small, that makes WiseMAC network traffic adaptive.

Sensor MAC [34] (S-MAC), a medium access control protocol, has periodic listen and sleep state for nodes as shown in figure 2.2. S-MAC has a unique feature of fragmentation in case of long messages and transmitting them in burst that allows access of medium for longer time to a node having more data.

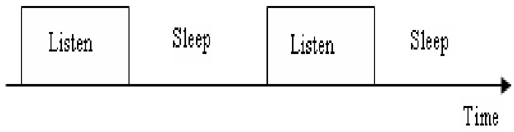


Figure 2.2: Periodic listen and sleep in S-MAC

Nodes have a periodic listen and sleep periods. Each node goes idle for a fix time and then again wakes up to listen the medium. For simplicity all nodes are assigned equal period of sleep and listen. All nodes are free to choose their schedule and share the information with their neighbors to synchronize their schedule with other nodes. Two neighboring nodes can have different sleep/wake schedule because they are can be synchronized to other nodes as depicted in figure 2.3. After synchronization, nodes are in a position to talk to other neighboring nodes despite having a different schedule. Nodes having synchronized schedules form a virtual cluster. In order to be perfectly synchronized with other nodes each node periodically broadcasts SYNC packets. The SYNC packets also help new nodes to join the network.

Figure 2.3: Neighboring nodes A and B have different schedules. They exchange with nodes C and D respectively

Multiple senders should contend for medium for sending to a receiver to avoid collision.

S-MAC implements classical RTS/CTS mechanism to avoid collision due to hidden terminal problem [35]. Nodes form a virtual cluster of synchronized nodes. But the synchronized nodes receive packets that are not intended to them causing overhearing and waste of energy. Overhearing is reduced by transiting nodes to sleep mode after hearing a RTS or CTS packet. Nodes which are immediate neighbors to both receiver and sender are put to sleep mode after receiving a CTS or RTS packet until the medium is not free as shown in figure 2.4.

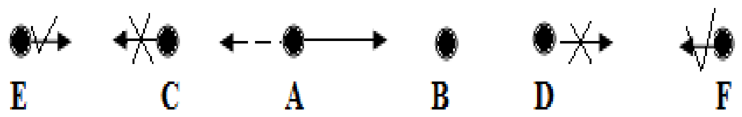


FIGURE 2.4: Who should sleep when node A is transmitting to B?

Long messages consume high energy in transmission and cause an energy problem if collision occurs that needs retransmission. With S-MAC long message is fragmented into small packets and then it sends all the packets in burst with single RTS/CTS. Hence the node reserves the medium for all fragmented packets.

Traffic-adaptive medium access protocol [36] (TRAMA) is an energy-efficient collision free channel access protocol for wireless sensor networks. In this protocol energy consumption is reduced by using collision free broadcast, multicast and unicast communication. A node, whenever it is not transmitting, is allowed to transit to a idle state which have low power consumption. TRAMA incorporate a collision free schedule which is based on the information of nodes that are one/two hopes away, ongoing time slot, and traffic information. Node scheduling is based on this traffic info which results in protocol being more adaptive for sensor network application. Protocol uses traffic information to schedule the nodes which makes the protocol more adaptive to sensor network applications.

TRAMA chooses random access mode to start and each node randomly selects a slot to transmit. Nodes are allowed to join the network during random access periods only. Random access periods facilitates the addition and deletion of nodes to the network. Time synchronization is also done in that period. TRAMA has three sub-protocols: Neighbor Protocol (NP), Schedule Exchange Protocol (SEP), and Adaptive Election Algorithm (AEA). NP gathers neighborhood information by exchanging small signaling packets to share information about network topology. SEP performs establishment and maintains traffic-based scheduling by allowing nodes to exchange information about traffic in their two-hop neighbors. AEA then constitutes a collision-free slot for a data transmission based on the information gathered by NP and SEP.

Flexible-Schedule-based TDMA Protocol [37] (FlexiTP) is designed for periodic data-gathering applications. FlexiTP does not unnecessarily hold the number of slots required for network. So in FlexiTP, nodes are free to join or leave the network any moment without affecting the performance of protocol which makes the protocol fault-tolerant.

FlexiTP consists of mainly four functions: route establishment (construction of data-gathering tree), deciding time slots of nodes, node synchronization and local repair for node failure. To manage the mentioned functions, two phases of FlexiTP are: initial network setup and data-gathering phase. Former is executed once and builds a data-gathering tree along with the scheduling the nodes. Nodes maintain the same schedules throughout the network lifetime. CSMA/CA is used in initial setup phase. Local topology, i.e. parent, children, descendants and first-level neighbors, is known to each node because of data-gathering tree. Forwarding-to-parent routing scheme [38] is applied over the data gathering tree.

Base station manages the time slot assignment phase by generating a time slot assignment token and passes it to the network. Depth-first-search technique is used for token passing. Depth-first-search scheme reduces buffering. Nodes inform the base station about the highest slot of network through GHS (Global Highest Slot) assignment. Multi-hop parent-children synchronization scheme [39] is performed for node synchronization.

Z-MAC [40], a hybrid MAC protocol, combines the strengths of CSMA and TDMA and avoids their weaknesses. Z-MAC applies CSMA and TDMA as demanded by the channel contention. For a low channel contention, Z-MAC uses CSMA while for high contention TDMA is applied. DRAND [41] is used for channel scheduling. Set-up phase of Z-MAC is applied once until there are significant changes in topology. Setup phase performs the operation of neighbor discovery, slot assignment, local frame exchange and global time synchronization.

Conclusion- MAC protocols for wireless sensor networks avoid/reduce the main sources of energy consumption of nodes like, collision, overhearing, idle listening, overemitting. Literature of MAC protocols suggests that schedule-based MAC approaches are more energy efficient. TDMA based approaches has a natural advantage of collision-free medium access. Additional collision avoidance or collision detection methods are required for CSMA based protocols to handle the collision possibilities.

2.2 Sleep/Wake-up Scheduling Algorithms for Wireless Sensor Networks

High density of nodes in sensor network provides better coverage and reliability to network but also produces a large amount of redundant data because a region is sensed by a number of nodes. Putting some redundant nodes to off mode while maintaining the coverage of network do not affect the overall performance of network. *Sleep/wake-up scheduling* protocols decides the on-duty and off-duty nodes based on some requirements of application while maintaining some redundancy [42]. Each time slice, nodes selected for on-duty perform the task of sensing on behalf of their redundant nodes. Sleep/wake-up scheduling protocols can be integrated with MAC protocol, but it is independent of the utilized MAC protocol because it sits above the MAC.

GAF [43] identifies and turns off the routing equivalent nodes thereby reducing the energy consumption of nodes. GAF divides the complete area in virtual grids by using location information of nodes as shown in figure 2.5. The virtual grids are such as nodes of two adjacent grids can communicate to each other directly, i.e. all nodes of grid A and adjacent grid B can can communicate with each other. So all the nodes in a grid are treated as equivalent for routing. Size of virtual grid (r) in GAF depends on the radio range of node (R).

$$r \leq \frac{R}{\sqrt{5}} \tag{2.1}$$

Nodes can be in one of the three states; *discovery, active, sleep.* Initial state of all nodes is discovery state. In discovery state all nodes have their radio on and exchange discovery message to find nodes in the same grids to form virtual grids. So to broadcast

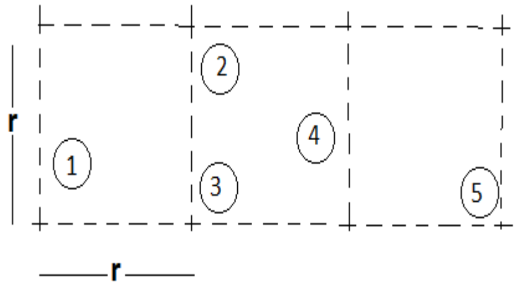


FIGURE 2.5: Virtual grids in GAF

discovery message, nodes transit to active state. Nodes periodically broadcast discovery message for Ta time duration. If a node either in discovery state or active state discovers an equivalent node of high rank to handle the routing, it will transit to sleep state. After Ts duration, node comes back to discovery state from sleep state. Change of state by nodes lead to load balance network.

In an event-driven scenario, most of the time nodes remain in monitor state to sense the environment. For most of the time as soon as an event occurs in a local region nodes start transmitting the data. The scheme **Sparse Topology and Energy Management** (STEM) [44] reduces the energy consumption in the monitoring state thereby increasing the network lifetime while ensuring a limited latency for transition to transfer state. Two different radios for wake-up signal and data packet transmission as shown in figure 2.6. Nodes only switch on the sensors and preprocessing circuitry while communication subsystem is turned off. When an even is followed, the main processor is waken for analyzing the data in more detail and transmission radio is waken up only if the data is of any interest. Node sends a beacon with the ID of the node it is trying to wake up. Each node then turns on periodically its radio for a short time. The target node detects the beacons and responds to the initiator node. All processing to wake up nodes is done by the radio operating in frequency band f1. Once a path is established between initiator and target node, both nodes initiate their radios that operate in frequency band f2 to transmit data packet.

Multi-parent scheme proposed in [45] takes multiple routes for transferring messages and wake-up schedules for various nodes. There are two communication paths in the network: forward direction and backward direction. The base station sends a message

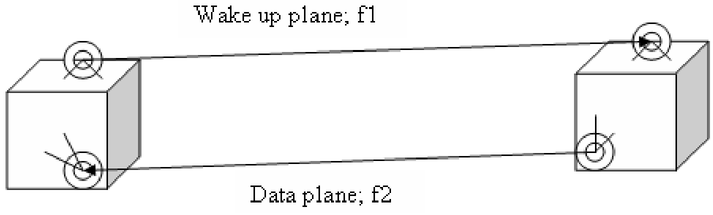


FIGURE 2.6: Sensor node with two radios operating on two frequencies

to one of the nodes in the network on forward direction path. In backward direction, a regular node communicates the message to the base station.

In a dense network, most of the nodes at highest levels have a number of neighbors which can communicate with many nodes that are a lowest level. Multiple-parent scheme has multiple parents and multiple paths with different wake-up scheduling for each node. Nodes in the network are divided into multiple disjoint groups such that at least one node from a group of parents can be assigned as parent. The base station, which is a special node, belongs to all the groups and should be in wake-up state all the time. If there are g groups, then nodes of a group follow the same wakeup pattern and sleep in the other (g-1) frames. Scheduling of nodes with single parent and multiple-parent with 2 groups is shown in figure 2.7.

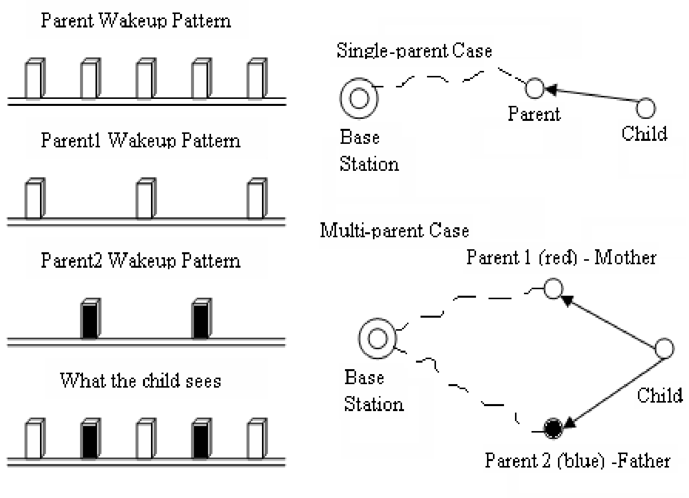


FIGURE 2.7: Multi-parent Methods

A **two tier scheduling** (TTS) algorithm proposed in [46] combines both connectivity preserving scheme [47, 48] and coverage preserving scheme [49, 50]. Nodes are selected for coverage tier and among these nodes a set of nodes is selected for connectivity tier. Nodes not selected for either coverage tier or connectivity tier are put into sleep mode. Nodes in coverage tier wake-up periodically and sense the environment and send the data to sink, while nodes for connectivity nodes are always active to relay the data of coverage tier node. Nodes of coverage tier which are not in connectivity tier may go to sleep state periodically. Connectivity- and coverage tier are rotated dynamically to load balancing.

A weighted greedy algorithm is used to select the nodes of coverage set, which are enough to detect all events in the entire sensing field. Residual energy of nodes is considered as weight in the greedy algorithm. A connected dominating set is selected from coverage set. Nodes having higher residual energy and degree of connectivity result in a higher chance of being a dominating node. Energy consumption of connectivity tier nodes may be higher than the energy consumption of the nodes of coverage tier. Other nodes are in sleep mode so have minimum energy consumption. Schedule of nodes with tiers is shown in figure 2.8. To balance the energy consumption, coverage and connectivity tier are updated periodically after every TR. In each round, current residual energy of nodes is used to calculate cost function. Therefore, there are fewer chances for a node having less energy to be selected for coverage tier.

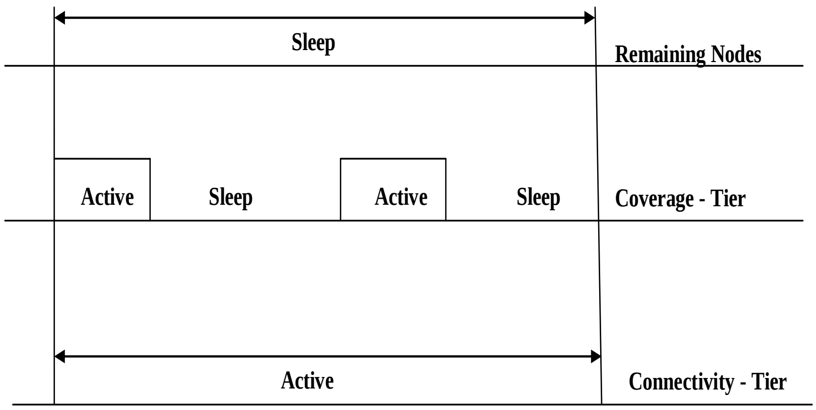


FIGURE 2.8: Sleep/wakeup Schedule of coverage-and connectivity- tier in TTS

Conclusion- Sleep scheduling algorithms periodically selects a set of nodes which works for the rest of other nodes. These algorithms conserve energy by exploiting dense deployment of nodes and reducing redundancy from the network.

2.3 Clustering Algorithms for Wireless Sensor Networks

Clustering is a key technique used to extend the lifetime of a sensor network by reducing energy consumption. Section 1.3 detailed the classification of clustering algorithms for wireless sensor networks. This section briefs the clustering algorithms proposed in the literature and divides the work in two categories namely - *Homogeneous Clustering Algorithms* and *Heterogeneous Clustering Algorithms.*

2.3.1 Homogeneous Clustering Algorithms

Low-Energy Adaptive Clustering Hierarchy(LEACH) [51] is a fully distributed clustering algorithm that selects cluster head randomly and assigns the role of cluster head to all nodes in rotation. LEACH protocol is considered as the revolutionary in the field of clustering algorithms for wireless sensor networks and also from its inception is the basis for most of the clustering algorithms for wireless sensor networks. Operation of LEACH is performed in rounds and each round has two phases: set-up phase and steady phase. In the former phase, sensor nodes self-organize in local clusters. Initially, each sensor node decides its role for that round, whether it will be a cluster head or a member node. Each node then selects a number randomly between 0 and 1. If the selected number is less than the threshold for the current round, node will be selected as cluster head. Threshold is calculated as:

$$T(n) = \begin{cases} \frac{P}{1-P\times(r mod \frac{1}{P})} & \text{if } n \in G \\ 0 & \text{otherwise} \end{cases} \tag{2.2}$$

where, P = desired percentage of cluster heads in network,

r = current round, and

G = set of nodes that have not been cluster heads in the last 1/P rounds.

For round = 0, all nodes are eligible for becoming cluster head and have equal probability of being cluster head. Each node is selected as cluster head once in 1/P rounds.

Nodes selected as cluster heads broadcast their status message to the network and wait for the response from other nodes. Nodes receive advertisement of all cluster heads and calculate received signal strength (RSS) of each advertisement message. Node selects cluster head with maximum RSS of advertisement message and sends request to join that node. Each node joins one cluster head. Cluster head nodes have complete information about clusters and each cluster head forms TDMA schedule for the cluster. Cluster head then broadcasts TDMA schedule to the cluster. Now, the set-up phase of LEACH is completed.

In steady phase, nodes send the sense information to the cluster head according to TDMA schedule. Node wakes-up as the allotted time slot in TDMA schedule arrives, sends data to cluster head and transits to sleep state to save energy. In LEACH, nodes are in sleep state most of the time, so nodes have low duty cycle. Data fusion or data aggregation process is applied to reduce collected data to some meaningful information. Cluster head sends aggregated data to base station. After a fixed time, nodes again go for set-up phase. Duration of steady phase is made longer than the set-up phase to avoid overhead of frequent clustering.

LEACH-C [52] is a centralized derivation of LEACH. Base station receives information regarding energy and location of a node, from the respective node. Base station then uses simulated annealing [53] to select cluster head, forms clusters, perform TDMA schedules and thereafter broadcasts the information to the network. Second phase, i.e., steady phase of LEACH-C is same as that of LEACH. Cluster head receives information from the nodes, aggregates it and forwards this aggregated information to base station.

LEACH-F [54] is an adaption of LEACH and LEACH-C but it has fixed cluster formation. As networks commence the operation, nodes communicate their energy and location information to base station. Depending on these information base station performs clustering and assigns nodes to different clusters ensuring efficient grouping. A list of cluster heads for each cluster is then decided by the base station. Base station also determines the TDMA schedules for each cluster and broadcasts this entire cluster information to the network. LEACF-F has fixed cluster formation because

- formation of cluster takes place only once
- cluster heads are selected from cluster head list
- TDMA schedule remains unaltered for the entire duration of network lifetime.

- set-up phase occurs just once.

Steady phase of LEACH-F, LEACH and LEACH-C are same.

In [55], author proposes an improvement over LEACH for energy conservation at nodes. To select the cluster head, proposed algorithm applies both LEACH approach and modified approach. If the remaining energy of network is greater than 50% of initial energy, LEACH approach is used to select the cluster head. If the residual energy is below 50% of initial energy, nodes with the highest remaining energy have high probability of being selected as cluster head. For cluster head joining, a cost function - equation 2.4 - is used which has the remaining energy of cluster head node and RSS of advertisement message of that cluster head node.

$$Cost(i) = Remianing\ Enegry\ CH(i) + RSS\ of\ CH(i) \tag{2.3}$$

Each node calculates the cost of each cluster head node and joins the cluster head with maximum cost value. In data transmission phase if a predefined condition, e.g. Does the temperature exceed 30°?, is satisfied by the sensed data only then the node transmits data to cluster head.

ESCAL [56] is based on LEACH. Any cluster head sends the gathered data to a nearby cluster head close to base station unlike LEACH where cluster head sends the data directly to base station. Therefore, a cluster head need not to transmit the data to a long distance resulting in significant energy conservation.

Enhanced Centralized LEACH [57] (ECLEACH) follows a centralized approach in which the base station is responsible for the selection of cluster heads. ECLEACH selects the cluster heads based on their residual energy, their distance to other nodes, and the residual energy of the other nodes. ECLEACH also keeps a minimum distance between every cluster heads and the next in order to have a better distribution of cluster heads over the network. The base station starts the selection process at the beginning of each round with calculating a threshold for each node according to equation 2.5. Then, the base station selects the node with the highest threshold to be the first cluster head. Then, it checks if the distance between the node with the highest threshold and the node with second highest threshold is greater than or equal to MDBECHAN. If it is, then the node with the second highest threshold is going to be the second cluster head. Otherwise, it checks the distance between the node with the highest threshold and the node with third highest threshold in the same way. The base station is not going to select two consecutive cluster heads if the distance between them is less than MDBECHAN. This is important to ensure that the cluster heads are properly distributed over the network. If the base station cannot meet the condition of MDBECHAN, then it needs to decrease the MDBECHAN value. When all cluster heads are selected, the base station broadcasts the list that contains the cluster heads in the current round to all nodes. Each member node sends its residual energy together with its sensed data to its cluster head which it turn forwards it together with the aggregated data to the BS.

$$T(n) = \frac{RE(n)}{\sum_{i=1}^{m} \frac{D(i,n)}{RE(i)}} \tag{2.4}$$

where, RE(n) is the residual energy of the node n,

m is the number of SNs in the network,

D(i, n) is the distance between node i and node n and is equal to 0 if i = n,

and RE(i) is the residual energy of node i.

Adaptive decentralized re-clustering protocol (ADRP) [58] is a base station assisted protocol which chooses a set of cluster heads for ongoing rounds and few upcoming rounds. Cluster heads and sets of cluster heads are selected according to the remaining energy of sensor nodes and average energy of each cluster. Operation of ADRP is shown in figure 2.9.

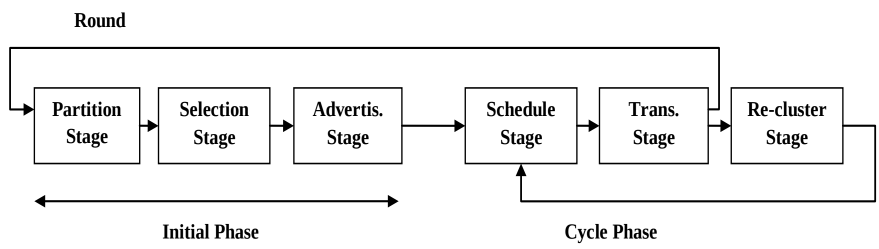


FIGURE 2.9: Operation of ADRP

In the initial phase, nodes communicate their location and energy information to the base station. Base station processes this information and divides the network into clusters and assigns one cluster head to each cluster, taking into account energy consumption approximation. Subsequently, base station decides cluster head nodes for upcoming rounds. Base station then broadcasts the cluster information, current cluster heads and next cluster heads.

In the cycle phase, cluster head broadcasts TDMA schedule. Cluster head receives data from nodes, aggregates it and transmit this aggregated data to base station. At the end of this phase, new cluster heads are selected from the set of next cluster heads if the list is not empty; otherwise initial phase is executed. Nodes avoid re-clustering for next few rounds as cluster heads are determined in advance, resulting in less energy consumption.

Energy Efficient Clustering Protocol The main goal of EECPL [59] is to distribute the energy load among all sensor nodes to minimize the energy consumption and maximize the network lifetime of wireless sensor networks. EECPL is a base station assisted centralized clustering approach. Each node sends its energy status and location to the base station. The base station uses this information to find the number of cluster heads and cluster senders as well as making sure that only nodes with enough energy participate in the cluster heads and cluster senders selection. The cluster head is responsible for creating and distributing the TDMA while cluster senders responsible for sending the aggregated data to the base station. When the clusters formed, the EECPL organizes the sensor nodes within cluster into a ring topology so that each sensor node receives data from a previous neighbor and transmits data to a next neighbor as shown in figure 2.10. Once the cluster heads and cluster senders are determined, the base station broadcast to all sensor nodes the information including cluster heads, cluster senders and sensor nodes IDs within ring.

Cluster head creates and distributes the TDMA schedule, which specifies the time slots allocated for each member of the cluster. Cluster heads create TDMA schedule telling each sensor node when it can receive and transmit the data. For gathering data in each frame, the cluster senders sense their environment, collect sensed data and transmit the data to the next neighbors clockwise. Each sensor node receives data from previous neighbor, aggregates with its own data, and transmits to the next neighbor on the ring.

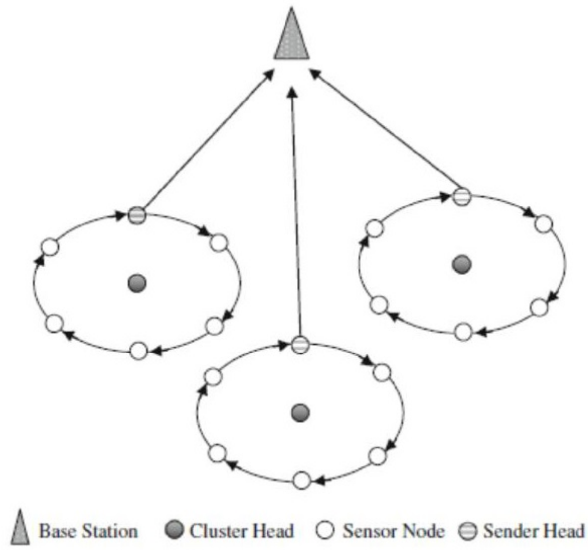


FIGURE 2.10: Clustering in EECPL

Upon receiving the aggregated data from previous neighbors, cluster senders transmit it to the base station.

A. Azim et al. in [60] addresses the fixed round problem in LEACH, LEACH-C and LEACH-F. Due to fixed round-time, LEACH and variants suffer from premature death of cluster head that causes loss in data packets. As the network proceeds, nodes consume energy and cluster heads no longer have enough energy to finish the long fixed round-time and the cluster loses information completely. Proposed solution implemented with LEACH-F calculates round-time by taking into account the remaining energy of cluster head. Low remaining energy cluster head nodes have small round-time and hence send the data to base station before completely going out of energy. A. Azim et al. extend their work [60] in [61] from LEACH-F to LEACH and LEACH-C (i.e. from fixed to dynamic clustering). Round-time is measured according to heuristic depending on the number of rounds.

Unequal Cluster-based Routing (UCR) [62] solution proposed by G. Chen et al. exposed the hot spot problem in clustering algorithms in which cluster heads are sending data to base station relying other cluster heads. In these clustering algorithms cluster heads near to base station consume more energy than the other because the have extra load of communication imposed by other cluster heads hence are more prone to early failure, leaving areas of the network uncovered and causing network partitions. UCR groups the nodes into clusters of unequal sizes. Cluster heads closer to the base station have smaller cluster sizes than those farther from the base station, thus they can preserve some energy for the inter-cluster data forwarding. For load balancing, a greedy geographic and energy-aware routing protocol is designed for the inter-cluster communication, which considers the tradeoff between the energy cost of relay paths and the residual energy of relay nodes.

Energy-balanced unequal clustering (EBUC) [63] applies PSO [64] to optimize the cluster formation to reduce intra-cluster communication distance and to form unequal-sized clusters. Clusters near to base station are of small size to conserve energy for inter-cluster relay of data packets as shown in figure 2.11.

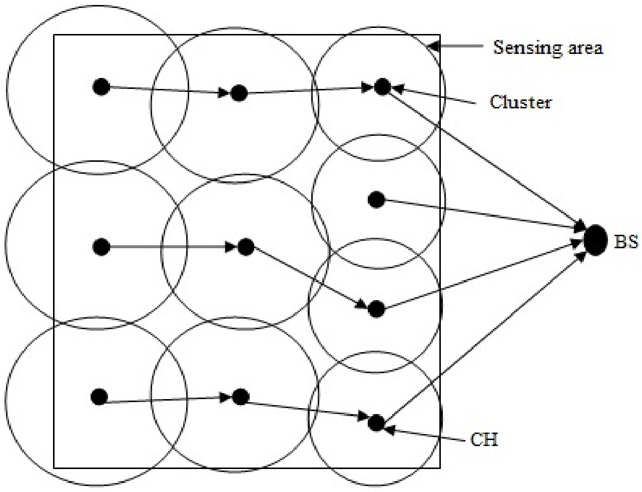


FIGURE 2.11: Cluster formation in EBUC

Operation of EBUC is performed by base station in a centralized manner. EBUC operates in rounds. In set-up phase of first round, nodes send their energy and location information to base station. Base station applies PSO with a cost function for having clusters of different size and with reduced inter-cluster communication. Base station broadcasts the cluster information to the network. In steady phase, nodes sense and send the information to cluster head R times in a round (value of R is predefined). Base station estimates the energy dissipation of nodes hence nodes are not suggested to send the information again.

In [65], author describes theoretical aspects of the clustering to minimize energy consumption by reducing communication distances of member nodes to cluster head. The balanced k-clustering problem is optimized by min-cost network flow. K is the number of cluster head nodes. In balanced k-clustering, clusters are balanced (in terms of sensor nodes) and one cluster head is chosen for each cluster. Proposed method by using min-cost network flow optimizes the total intra-cluster communication distance.

Hierarchical cluster-based routing (HCR) [66] is a base station controlled algorithm. Nodes communicate their energy and location information to base station. Base station computes the best possible cluster heads and broadcasts it to the network. Nodes join the nearest cluster head node. Cluster head node then selects a set of associated head nodes. Associated cluster head nodes do not participate in sending the data to cluster head. Only one head among cluster head and associated head nodes will be active at a time. Cluster head node and associated head nodes receive the data from member nodes and send the data to base station along with the energy information of nodes.

In [67], author presents a distributed clustering algorithm with multi-level hierarchy. Single level hierarchy of cluster head is extended to increase the level of cluster heads. Each sensor node can become a cluster head with probability, p, and announces to the network about its cluster head status. These cluster heads are known as volunteer cluster heads. The advertisement of volunteer cluster head is forwarded only to k hops in the network. A non-cluster head node receives these advertisements and selects the cluster head that is nearest to it. If an advertisement message is not received by the sensor node within a time period, t, it will announce itself as cluster head and broadcasts status message. These nodes are called forced cluster head nodes.

Power aware clustered TDMA [68] (PACT) is an energy efficient TDMA protocol in which the duty cycle of nodes is adaptive to the user traffic. PACT combines the passive clustering [69] and TDMA to reduce the power consumption. Passive clustering eliminates the need of control messages and also effectively solves the isolation problem. Passive clustering allows nodes to take turn to become the communication backbone nodes. PACT extends passive clustering by taking into account the energy information. A new state of node called Low Energy State (LES) is also introduced with the other state of passive clustering (e.g. cluster head, gateway and ordinary node). PACT changes the status of a node in a distributed manner based on its battery energy level. Status of a cluster head node or gateway node is changed to LES when the energy level of that node drops below a certain threshold.

The number of active gateways between neighboring cluster head are limited. Each node has information of cluster head IDs detailed by its neighbors. This information is used to limit active gateways between neighbor cluster heads. PACT uses two hop neighbor scheduling information to avoid conflicts and collision. Protocols adapt duty cycle of the nodes to user traffic where radios are scheduled to be turned off during inactive periods.

Energy-efficient protocol with static clustering (EEPSC) [70] does the partitioning of network in clusters once. Once clusters are formed, selection of cluster head nodes and temporary cluster head nodes is done dynamically. For clustering of network, base station broadcasts k-1 different messages of different power, where k is desired number of clusters. Nodes in the network receive the message and send back the JOIN-REQ message to base station with message ID. Remaining nodes set their cluster ID to k. Partition of network in clusters is shown in figure 2.12.

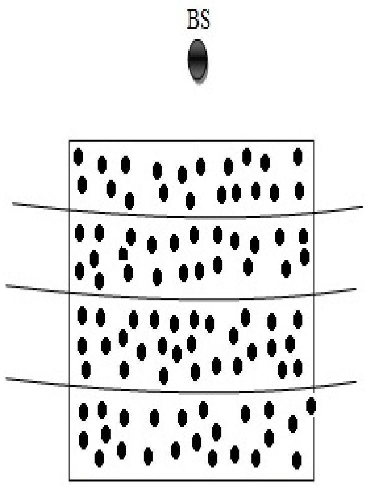


FIGURE 2.12: Cluster formation in EEPSC

Base station randomly selects a temporary cluster head for each cluster and advertises it to the network. Base station also broadcasts TDMA schedule of each cluster. TDMA schedule is complied once. Temporary cluster head receives information from nodes of cluster and selects cluster head with maximum remaining energy and temporary cluster head with minimum remaining energy for the next round. Nodes send the data to cluster head. Cluster head forwards collected data to base station after aggregation.

Clustering Protocol (CP) [71] is based on the Covering Problem. CP ensures that cluster head is at the center of hexagonal cluster and all member nodes are within the transmission range of cluster head node. Base station always acts as Initiator every time because base station has unlimited energy. Initially all nodes set themselves as un-clustered. Initiator sets some random orientation to define its hexagonal cluster. Initiator broadcasts its cluster head advertisement with limitation of 2-hops.

If an un-clustered node A receives CHA directly from a cluster head node, it will set that node as its cluster head. If an un-clustered node X receives CHA via some other node, it calculates orientation, position and distance to the center of cluster and sets timer to t = f(d). If node X does not receive any CHA directly before time expires, it will announce itself as cluster head. If it receives CHA directly, it will join that cluster and nullify timer. Value of f(d) depends on the density of nodes and the time required to process/transmit/receive a message. Reconfiguration of clusters is done to adapt network dynamics.

Energy-efficient data aggregation protocol based on static clustering (EDASC) [72] is a static clustering protocol based on Hausdorff Distance. If the network is partitioned into M clusters G1,G2,...,GM, and the Hausdroff distance is d, energy node in G_i must be within a distance d from some node in G_j and vice versa. The worst case of two neighboring clustering is shown in figure 2.13.

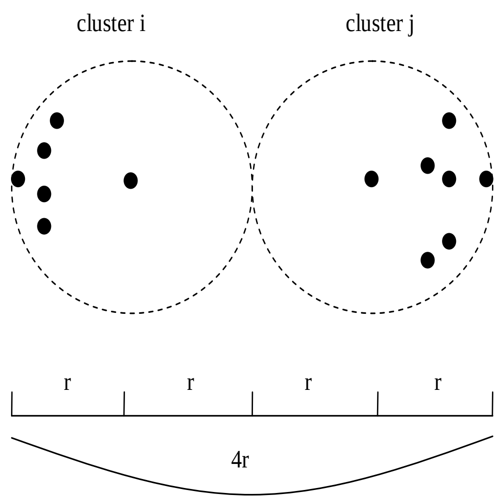


FIGURE 2.13: Two neighboring clusters in EDASC (Worst Case)

The worst coverage distance between two closet nodes is 2r. C and D are cluster heads of respective clusters, and then maximum distance for inter-cluster is 4r. Base station selects a random initiator. Initiator broadcasts message and waits for JOIN-REQ message.

A node will join that cluster head if two conditions are satisfied.

1. The Hausdorff distance between node and cluster must be smaller than r.
2. If the node is admitted, the Hausdorff distance between two neighboring clusters must be no longer than 3r.

If the conditions are not satisfied, the node will announce itself as cluster head. After cluster head selection, data aggregation (DA) tree of cluster head nodes is constructed. Cluster head near to base station is the root of the tree. If there are some changes in cluster head topology, base station reconstructs DA tree and broadcasts them to all nodes.

QoS-based adaptive clustering(QAC) [73] is a local-centralized, dual cluster head selection algorithm that increases the quality of service with an increase in the lifetime of wireless sensor networks. QAC assumes that interim cluster head nodes are deployed with other nodes and base station activates them after deployment. In the set-up phase of first round, initial interim cluster heads elect themselves as master cluster heads and announce to the network. But in other rounds, master cluster head nodes are selected according to the information collected in previous round by interim cluster heads.

Nodes join the nearest cluster head by comparing the received signal strength (RSS) of advertisement messages. Let there may be K nodes in a cluster. If $Y \leq K$, there is no need of slave cluster and cluster set-up phase is completed. But, if $Y > K$, there will be requirement of slave cluster. Master cluster head will select a slave cluster head node that has the smallest RSS and energy above average energy of cluster. Nodes of this cluster will choose cluster head again among master and slave cluster head. Cluster formation is shown in figure 2.14. After master and slave cluster formation cluster formation phase is completed.

Each master and slave cluster head node creates TDMA schedule and broadcasts it to the respective cluster. If a master cluster head node goes out of energy during the operation, slave cluster head takes care of complete cluster, and vice-versa. After data transfer phase, master cluster head nodes change their role to interim cluster heads after Tw time, and set-up phase starts again.

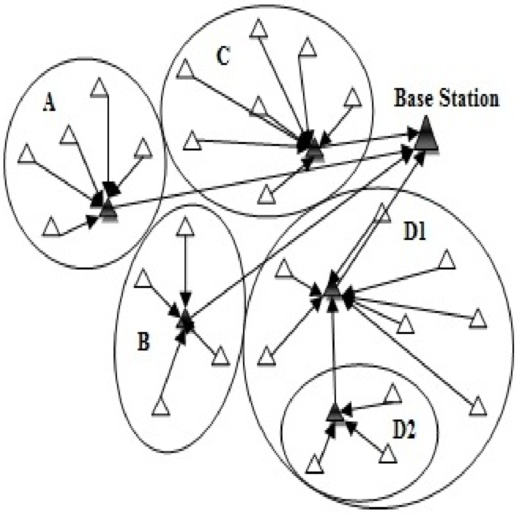


FIGURE 2.14: Cluster formation in QAC

In [74], author uses Fuzzy C-Means (FCM) [75] clustering approach to find optimal number and location of cluster heads in wireless sensor network to prolong the network lifetime. Euclidian distance is used by the approach to partition the sensor network in clusters and Xie and Benis (XB) index is used as validity measure of clusters.

Genetic Algorithm [76, 77] is applied in [78, 79, 80, 81, 82] to find the best suitable cluster head nodes for each round.

2.3.2 Heterogeneous Clustering Algorithms

Stable Election Protocol (SEP) [83] enhances the stable region of network lifetime because it is capable of handling the heterogeneous nodes. Nodes are divided into two groups: advance nodes (high energy nodes) and normal nodes(relatively low energy nodes). These two groups have different probabilities of selection as a cluster head.

$$p_{nrm} = \frac{p_{opt}}{1 + \alpha m} \tag{2.5}$$

$$p_{adv} = \frac{p_{opt}}{1 + \alpha m} \times (1 + \alpha) \tag{2.6}$$

,where P_{opt} = optimal probability to become a cluster head,

m = fraction of advanced nodes of total nodes consider as advance nodes, and

α = additional energy factor for advance node.

Epoch, the total number of rounds when each node is selected as cluster head once, of the clustering algorithm is increased so that advance nodes can act as cluster head for more than one round. Epoch is further divided into sub-epochs. Each normal node becomes cluster head once in an epoch while every advance node will be a cluster head once in each sub-epoch. [84, 85] have proposed extension to SEP.

Distributed Energy-Efficient Clustering (DEEC) [86] algorithm highlights the heterogeneity of sensor nodes to lengthen sensor nodes lifetime which is dictated by Yarvis et al. in [26]. DEEC suggests that nodes with high remaining energy should be the cluster head in comparison to the nodes with low residual energy. Probability of being the cluster head is different for nodes according to their state as in [83]. In DEEC, probability also depends on the remaining node energy and average network energy.

$$p_{adv} = \frac{p_{opt}(1+\alpha)E_i(r)}{(1+\alpha m)E_{avg}(r)} \tag{2.7}$$

,where E_i(r) = remaining energy of node i for round r,

E_{avg} = average energy of the network for round r.

Advanced nodes have high energy so have high probability of the selection of cluster head.

Energy efficient heterogeneous clustering (EEHC) [87, 88] works with three level of heterogeneity: 1. Super nodes (having highest energy), 2. advance nodes (have less energy then super nodes but still higher), 3. normal nodes. Probability for a node to be selected as a cluster head depends on node energy hance super nodes have significantly high probability to be selected as a cluster head than advance nodes and normal nodes. EDDEEC in [89] proposes an extension to DEEC and EEHC by considering switching behaviour of super and advance node to normal node when the residual energy is less than $T_{absolute}$.

Weight Based Clustering for Heterogeneous Networks(WBCHN) [90] utilizes heterogeneity in sensor networks to extend its lifetime and throughput by electing better cluster heads in a proficient manner. The main objective of WBCHN is to enhance the stability period by electing sensor nodes with higher residual energy as cluster head, to elect cluster head in a deterministic manner based on residual energy and to prevent the election of low energy sensor nodes. Proposed WBCHN solution takes into account (i) the residual energy, number of live neighbors of a sensor node and (iii) its distance from the Base Station. A sensor node elects itself as a cluster head if it has residual energy greater than the average residual energy amongst its neighborhood. To predict the number of live neighbors accurately, energy single sensor node broadcasts a *I am alive* message to all its neighbors at the end of each round. WBCHN hypothesizes that election of both the advance nodes, lie within the same neighborhood, enhances network lifetime.

Cluster Head Re-election Protocol (CRP)[91] solution exploits heterogeneity of sensor nodes to select cluster head with higher residual energy for each cluster in each round. Cluster head selection is done in two phases. In the first phase, nodes are selected for the role of cluster head as done in [51, 83] depending upon the distance of base station from the network. Cluster heads elected within this phase are called tentative cluster heads. The selected cluster heads broadcast their advertisement about cluster head status. Nodes join nearest cluster head according to Received-Signal-Strength of the advertisement message. Each member node sends a cluster join message which also has the information about the residual energy of node. In the initial phase, a sensor node can be selected having less residual energy than the any of the member nodes. So, the re-election phase select a new cluster having highest residual energy among the cluster. The information of new cluster head is broadcasted along with the TDMA schedule of cluster. This corrective measure leads to efficient usage of energy within sensor nodes and prolongs network stability.

D. Kumar in [92] presents and evaluates two distributed clustering algorithms viz. *single-hop energy-efficient clustering protocol* (S-EECP) and *multi-hop energy-efficient clustering protocol* (M-EECP) for heterogeneous wireless sensor networks. Three types of sensor nodes differing in initial energy are considered for heterogeneous environment. The proposed solution has cluster selection in such a way that the network stuck problem after certain number of rounds while there are still alive nodes is solved. The threshold for cluster head selection procedure is expanded by a factor which increases the threshold for any node that has not been selected as cluster head for the last 1/p rounds. In S-EECP, cluster heads send collected data directly to the base station while in M-EECP, cluster heads send collected data via some other cluster heads.

Energyaware routing protocol (ERP) [93] exploits difference in energy of nodes for cluster head selection. ERP reduces the energy consumption by area coverage problem. If a node has higher energy than average residual energy of neighboring nodes in the cluster range then it has better chances of being a cluster node. As a round starts, nodes update their neighbour table by communicating information messages to each other and then each node calculates the residual energy of each node in cluster range. Then broadcasting delay time is evaluated, based on residual energy and average energy of nodes in cluster range and thus cluster head selection is achieved.

Eventually nodes that have higher residual energy are selected as cluster heads. Normal nodes join the nearest cluster head node as members. Solving the coverage problem is demanded by the application. most of the nodes remain in sleep state and very few nodes are in active state thereby significantly improving the energy efficiency. Each cluster head calculates a weight according to the equation 2.9,

$$w_i = \frac{D(RSS_i) \times E_a}{D(RSS_{max}) \times E_{residual}} \tag{2.8}$$

,where RSS_i is received signal strength for a signal broadcasted by the base station of node i,

RSS_{max} is maximum received signal strength,

D( ) is the function that determines distance.

A cluster head node selects parent that has maximum weight in its range. In data transmission phase, nodes send the data to respective cluster head nodes. Cluster head nodes do aggregation and send the data to its parent node or base station.

Enrique J. Duarte-Melo et al. in [94] proposed a heterogeneous clustering approach which categorized nodes as overlay sensor nodes with more processing capability and communication capability in addition to having more energy, and normal sensors simple sensor nodes (without extra resources). Overlay nodes are also deployed randomly along with the normal nodes. Number of overlay nodes in the network is R.q (where R,q>1). On average only q overlay sensors are active at any round, i.e., on average there are q clusters. At the start of a round, each overlay sensor dynamically deciding whether it will be cluster head in the current round. If an overlay sensor decides not to be a cluster head for the current round, it goes to sleep for current round. An overlay sensor is active once and only once every R rounds. Normal nodes select the nearest active overlay nodes as cluster head. Proposed solution allows only overlay nodes for the role of cluster head consequently load balances the network.

N. G. Praveena et al. in [95] proposed a clustering approach with a single-hop communication based upon link correlation. The heterogeneous nodes are deployed which act as a cluster head and communicate with the base station. The proposed solution formed, using link correlation, a level-k cluster hierarchy (i.e. multi-level) with single-hop communication. The level-k clusters are the higher level in the hierarchy with $\{k-1, k-2, 1\}$ denoting the hierarchy of sub-clusters in the subsequent level. The nodes in each level act as cluster heads for its corresponding sub-level nodes and level-1 nodes act as leaf nodes. The energy level is considered while forming a cluster hierarchy, such that nodes of a higher hierarchy level have more energy than lower level nodes.

Each node in the network is connected in a single-hop communication to its corresponding cluster head in the above hierarchy level using link correlation. The level-k cluster heads form the bottleneck zone of the sink which has heavy traffic flow. This results in faster depletion of its energy reducing the network life time. To overcome it, the heterogeneous nodes are adopted as level-k cluster head, since it have more energy compared to normal nodes. After establishing a level-k cluster hierarchy using link correlation, the level-k cluster head forms a TDMA time slot for its corresponding a level-(k-1) cluster head, while the level-(k-1) cluster head forms a TDMA time slot for a level-(k-2) cluster head and it is followed for all the sub-clusters. The TDMA time slot adopted between cluster heads in subsequent levels of hierarchy helps to remove the data collision and cut-back the data aggregation (DA) time. A level-3 hierarchy was generated and evaluated for performance evaluation.

B. A. Attea et al. in [96] propose an evolutionary based routing protocol for clustered heterogeneous wireless sensor networks. Fitness function of proposed EA based approach consists of cluster's cohesion or scatter(intra-distance) and cluster separation (inter-cluster).

2.3.3 Conclusion

Clustering algorithms proposed for wireless sensor networks provide energy efficiency along with scalability, fault-tolerance, and many others features. Clustering algorithms conserve energy of nodes by avoiding the main sources of energy consumption in MAC. Most of these algorithms use TDMA within cluster and different codes to avoid intra and inter cluster collision respectively. Communication distances between the nodes and far located base station is avoided that reduces energy consumption of nodes - only cluster head nodes are communicating to base station. Re-clustering of nodes after certain periods make these algorithm obvious choice for fault-tolerance. Failure of few nodes is not affecting the overall performance of network.

In these algorithms, the selection of cluster heads is considered one of the main issues which are in charge of creating clusters and controlling member nodes, and a proper selection of cluster heads leads to reducing the energy consumption and prolonging the network lifetime. Literature also suggests that heterogeneous nodes are proficient for the role of cluster head as cluster head node has extra responsibilities than the other nodes consequently consumes more energy.

The compelling challenges for clustering algorithms are how to select cluster heads, location of cluster head in the cluster for intra-cluster and inter-cluster communication distances, and how to determine the optimal frequency for cluster head rotation in order to maximize the network lifetime.

Chapter 3

Network Model and Assumptions

This chapter describes the general network assumptions and radio model used for most of our work, simulators and frameworks used to do experiments on proposed solutions, and performance metrics for evaluating and comparing our proposed solutions against existing protocols.

3.1 General Network Assumptions

For the work of this thesis, the following network assumptions are considered:

- A single base station is situated outside the periphery of deployed area.
- Sensor nodes are deployed randomly all over the sensing area.
- Sensor nodes are not mobile, i.e. sensor nodes do not change their locations after deployment.
- In most of the work, sensor nodes are homogeneous with respect to initial energy and hardware specification. Chapter 5 considers heterogeneity of nodes. Nodes with same hardware specification but with different energy levels are considered.
- The number of optimal cluster head nodes is fixed at 5% [54].

3.2 Radio Propagation Model

Performance of wireless communication systems is limited by radio channel because unlike wired channel, radio channels are random in nature [97]. Transmission path between two devices can be a simple line-of-sight or it can be obstructed by many obstacles like wall, building, and mountain. Among several radio propagation models, there are three basic and prominent models: free space, two-ray ground reflection, and shadowing model. Work of this thesis considers free space and two-ray ground reflection model in the context of the distance between transmitter and receiver.

As described in [54], a predecided cross-over distance ($d_{crossover}$) is considered for switching from free space model to two-ray ground reflection model and vice versa. If the transmitter and receiver are apart by a distance less than $d_{crossover}$, free space model is considered else two-ray ground reflection model is taken. Value for $d_{crossover}$ is set to 87m.

Free space model is considered if line of sight path is available between transmitter and receiver (i.e. when transmitter and receiver are in vicinity to each other). In free space mode, received power of a signal received by an antenna is given by:

$$P_r(d) = \frac{P_t G_t G_r \lambda^2}{(4\pi d)^2 L} \tag{3.1}$$

where, P_t is the transmit power,

G_t = gain of the transmitting antenna,

G_r = gain of the receiving antenna,

λ = wavelength of the carrier signal,

d = distance between transmitter and receiver, and

L = system loss factor not related to propagation.

Geometric optics forms the basis for two-ray hround reflection model and hence takes into account both direct and reflected paths. In two-ray ground reflection model, received power of a signal received by an antenna is given by:

$$P_r(d) = \frac{P_t G_t G_r h_t^2 h_r^2}{d^4} \tag{3.2}$$

where, P_t is the transmit power,

G_t = gain of the transmitting antenna,

G_r = gain of the receiving antenna,

h_r = height of the receiving antenna above the ground,

h_t = height of the transmitting antenna above the ground, and

d = distance between transmitter and receiver.

For all simulation setups, all nodes are considered to have similar transmission power. Omni-directional antennas are considered with the following parameters: $h_t = h_r =$ 1.5m, $G_t = G_r = 1$, $\lambda = 0.328$ m, no system loss (L = 1) and 914 MHz radios.

3.3 Energy Model

Author in [97] suggests that the power attenuation of a signal depends on the separation between transmitter and receiver. Propagation losses are inversely proportional to d^2 if they are close to each other and, propagation losses are inversely proportional to d^4 for long distances. Energy is consumed in both transmitting and receiving of signal as depicted in figure 3.1.

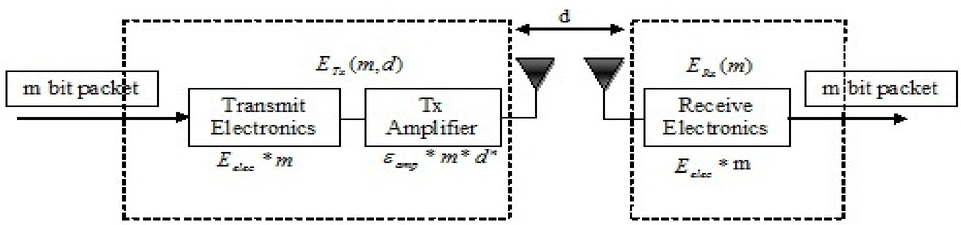


FIGURE 3.1: Radio Model

Energy consumed by the transmitter for transmitting an m-bit message to a receiver located at a distance d is given by:

$$E_{TX}(m,d) = E_{TX-elec}(m) + E_{TX-amp}(m,d) \tag{3.3}$$

$$E_{TX}(m,d) = \begin{cases} mE_{elec} + m\epsilon_{fs}d^2 & d < d_{crossover} \\ mE_{elec} + m\epsilon_{mp}d^4 & d \geq d_{crossover} \end{cases} \tag{3.4}$$

while the energy consumed by the receiver in receiving that message:

$$E_{RX}(m) = mE_{elec} \tag{3.5}$$

For considered network model with 1Mbps bandwidth, we set

$d_{crossover}$ = 87m,

E_{elec} = 50nJ/bit,

ϵ_{fs-amp} = 10pJ/bit/m^2, and

$\epsilon_{two-ray-amp}$ = 0.0013pJ/bit/m^4.

3.4 Experimental Set-up

In the literature, various methods are proposed for nodes deployment in the field like, non-uniform Regular Hexagonal Cell Architecture [98], Straight Line [99], Random [100], and many more, and hence for generation of network models. For the present work of thesis, network topologies of various size and number of nodes with random deployment are generated. Most of the work consider three network topologies of 50×50 m^2 with 50 nodes, 100×100 m^2 with 100 nodes and 150×150 m^2 with 200 nodes, unless otherwise stated. Random deployment of 100×100 m^2 with 100 is shown in figure 3.2. The base station is located outside the deployed area at 75 meters from the so called boundary of network. Each member node has data to send and each data message is 500 bytes long and packet header for each type of packet is 25 bytes long. The energy for aggregating data is set to 5nJ/bit/signal. The various parameters for the simulation are summarized in table 3.1.

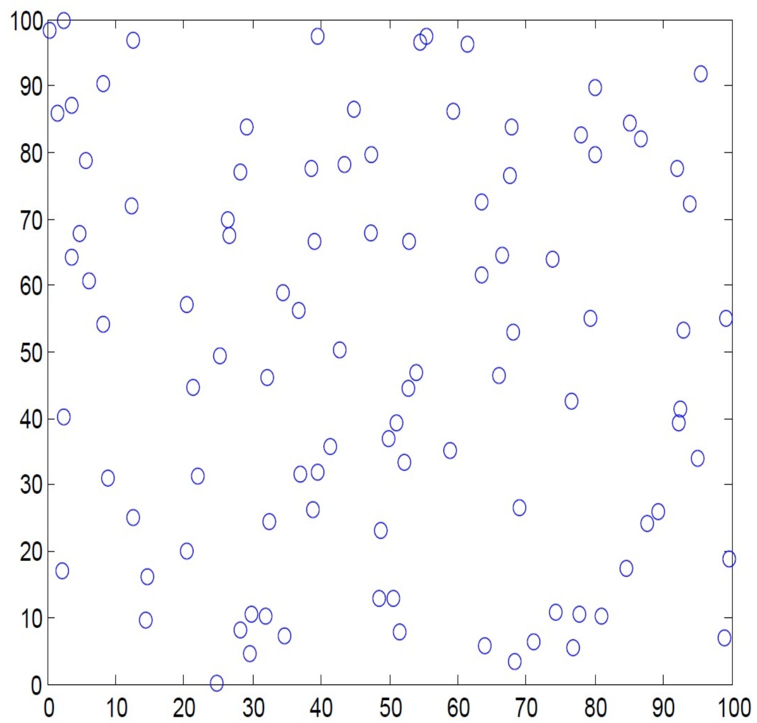


FIGURE 3.2: Deployed Network (Base station located outside the network area is not shown)

TABLE 3.1: Simulation Parameters and Values

Parameters	**Values**
Network Area	50×50 m^2, 100×100 m^2, 150×150 m^2
Number of Nodes	50, 100, 200
$d_{crossover}$	87m
E_{elec}	50nJ/bit
ϵ_{fs-amp}	10pJ/bit/m^2
$\epsilon_{two-ray-amp}$	0.0013pJ/bit/m^4
Header Packet Size	25 Bytes
Data Packet Size	500 Bytes
Bandwidth	1 Mbps

3.5 Simulators

For the performance evaluation of proposed solutions, we used ns-2 network simulator (version 2.27) [101] among other various simulators like NCTUns [102], Mannasim [103], OMNET++ [104], OPNET [105] and GloMoSim [106]. NS-2 was introduced in 1989 and since then it has continuously gained interest from research persons of every community like academia, industry and government [107]. Now, NS-2 is considered as the most utilized simulator for networks. NS-2 is an event driven simulator, i.e., it is initialized and run by a set of events.

Simulation of a network is carried forward from one event to another. C++ and Object-oriented Tool Command Language (OTcl) are the twp languages used by the NS-2. Internal mechanism of NS-2 is built in C++ while simulation is set by OTcl, i.e. OTcl works at frontend while C++ works at backend. TclCL links together C++ and OTcl. There are large numbers of built-in C++ object which are used to set up a simulation using Tcl simulation script. NS-2 supports both wired and wireless communication for simulation.

μAMPS extension [108] for NS-2 provides the required functionality for clustered wireless sensor network architecture and was the part of development of LEACH protocol. That μAMPS extension is implemented with NS-2. The extension includes MAC protocols, energy dissipation models for computation and communication and the protocol architecture of LEACH, LEACH-C, LEACH-F, MTE routing and static clustering. The extension supports change in these protocols. Simulation parameters, network topology with other network parameters are input for simulations. To handle the random nature of NS-2 and distributed characteristics of protocols, 30 different runs of simulation are carried and average of all these runs is taken for final results and comparisons, unless otherwise stated.

We also used MATLAB (MATrix LABoratory) [109] for our proposed solutions. MATLAB consists of hundreds of in-built functions to provide an interactive environment for researchers. Array is the fundamental data type while matrix is the main building block in MATLAB. It also provides platform to integrate with C and C++.

C and C++ platforms are used for the validation of proposed algorithms.

3.6 Performance Metrics

In this section, we describe the different metrics that we have been using in the present thesis work for performance evaluation for energy efficiency of proposed solutions and comparing these solutions with other important existing algorithms. Following energy efficiency metrics are taken into account:

1. **Network Lifetime:** It is a critical measure for energy efficiency of a protocol. It is the interval from the time of deployment of the networks until it is capable to fulfill the requirements of assigned task [7]. Efficiency of protocol is measured by the longevity of network lifetime. Depending on the application it broadly categories in the following metrics:

   - **First Node Death:** It is the time from the positioning of the nodes in the network until the first node stops responding [7]. The metric is important for applications that require high accuracy of information like health monitoring.
   - **50% Node Death:** It is the time from the positioning of the nodes in the network until the half nodes stops responding [7]. The metric is useful for applications like temperature monitoring.
   - **Last Node Death:** It is the time until network is able to function and transmit some information [7]. Optimal number of cluster head set at 5% so we are considering 95% node death as the last point of getting information, unless otherwise stated.

2. **Node Death Rate:** It is a measure of distribution of alive nodes over the simulation or the number of nodes which are dead over network lifetime. The region of the node death rate is divided into two parts: *stable region* and *unstable region*. Stable region represents presence of all nodes in the field hence it should be extended by the proposed solution. Remaining region is unstable region which should also be extended. A lower node death rate is achieved by a better load balanced network.

3. **Data Units Received at Base Station:** It is total number of data units successfully received at base station. The metric plays a vital role for data gathering application. Network lifetime directly effects the number of data units received at base station as longer the network lifetime more is the time duration available for sensors to sense and send the data to base station.

4. **Energy Consumption Rate:** It states the energy consumption of the whole network over the simulation time. It shows the energy efficiency of the algorithm over time. Energy consumption rate rely on various factors like lower the node death rate lower is the energy consumption rate.

The basis of selecting these metrics is that various proposed algorithms in literature [51, 52, 54, 57, 58, 59, 60, 62, 83, 85, 86, 87, 92, 95] have been elevated towards these metrics for performance evaluation and comparison. Energy efficiency is the sum up of above mentioned metrics [7]. It is obvious, the longer network lifetime are, the better does a network perform. Longer network lifetime of network solely cannot ensure the better energy efficiency of a solution. Increased network lifetime along with well load balancing and data gathering at base station define the energy efficient performance of algorithm [7]. An increased stable region of node death rate and low energy consumption is result of a well load balanced network. An increased stable region and longer network lifetime provides more sensor nodes for long time so there is increase in data collected at base station. To sum up, the improved performance for these metrics are result of better energy efficiency aspect of protocol architecture.

Chapter 4

Proposed Clustering Algorithm to Load Balance the Wireless Sensor Networks

The round-time and cluster heads selection of a clustering algorithm are very important to load balanced a wireless sensor networks. In this chapter, we propose a adaptive round-time method which dynamically compute the round-time based on the number of alive nodes further, a cluster head selection method has also been proposed which divides the network area into two parts- *Border Area* and *Inner Area*. The border area covers the nodes lying near the boundary of network area and rest of nodes are part of inner area. In the selection of cluster heads, only inner area nodes can participate for the role of cluster head. The performance of proposed methods has been evaluated on three different network topologies. The results show that the network lifetime and the data units received at base station have increased for adaptive round-time method as compared to the existing fixed round-time approach. The performance of the proposed cluster head selection has been found better than the traditional clustering approach, LECAH, in terms of network lifetime and data units received.

4.1 Introduction

One of the main challenges for clustering algorithms is that what should be the length of round-time. Performance of clustering algorithm rely to a great extent on the length of round-time [20]. If a clustering algorithm uses a long round-time then that algorithm reduces the energy consumption by reducing frequency of re-clustering but energy of cluster heads is depleted more compared to member nodes which results in a load unbalanced wireless sensor networks. If in this scenario, a low energy node gets selected as cluster head then that will falter before it completes the round and the cluster data will not be received by the base station. Though short duration of round-time does not drain much energy of cluster head but it over burden the network by the overhead of frequent re-clustering. As the re-clustering consumes energy of sensor nodes in exchanging control messages. This energy consumption grows much higher for large sized network. Clustering algorithms should take care of trad-off of long and short round-time for re-clustering [54].

Most of the existing clustering algorithms for wireless sensor networks use a fixed round-time. Fixed round-time poses a real challenge to the efficient operation of clustering algorithms. Because these algorithms calculate the round-time based upon the initial values of network parameters and use the same throughout network lifetime. Fixed round-time happens to be very large for energy deficient sensor nodes of the progressed network which was fair for initial deployment of network. Hence, the length of round-time should be adaptive to network dynamics [60].

Cluster head selection is significant issue for performance of the clustering algorithms. From the point of view of a clustering algorithm, the communication in a network can be labelled as: (1) *intra-cluster communication* - communication between member nodes and cluster head node, (2) *inter-cluster communication* - communication between cluster head nodes and base station (directly or via relay cluster heads) [24]. As all sensor nodes are involved in intra-cluster communication, it ensues greater energy consumption of network as compared to inter-cluster communication because it involves only cluster heads and base station. Position of cluster head node in the cluster is decisive for the performance of cluster. Clusters with cluster head node positioned near to the center have less intra-cluster communication distance while clusters with heads away from the center have high intra-cluster communication distance, therefor earlier clusters are more energy efficient than later clusters. Many of the existing clustering algorithm give equal importance to all the nodes for being selected as cluster head. For these algorithm, it might be possible that a node lying very near to the boundary of the network area may get selected as cluster head. It results in large intra-cluster distance for that cluster. Such clusters may leads to poor performance of the clustering algorithm. (It has been explained in Section 4.4 with the help of an example.)

The work of this chapter presents a clustering approach which dynamically calculates round-time that depends on the number of alive nodes in the network. Proposed method also takes care of trade-off of long and short round-time. Further, the proposed solution divides the deployed area in two regions: Border area and Inner area. Only inner area nodes are competent to play the role of cluster head while border area nodes are always member nodes. Remaining chapter is divided into several sections as follows: section 4.2 presents proposed solution of network adaptive round-time. Performance of network adaptive round-time solution is examined in section 4.3. Section 4.4 narrates proposed solution for cluster head selection which is anatomized in section 4.5. Section 4.6 concludes the work.

4.2 Network Adaptive Round-Time

Round-time is conclusive for the performance of clustering algorithms. Round-time is segregated into several frames and which are further divided into time slots. A sensor node is then assigned one time slot per frame. So, round-time (t_{round}) can be calculated as [54]:

$$t_{round} = X_{frames/round} \times t_{frame} \tag{4.1}$$

where X = number of frames per round and,

t_{frame} = length of one frame (in time).

Let there are N nodes distributed in M×M area with base station is situated outside the deployed area. The number of optimal clusters per round is k and there are N/k member nodes in each cluster. Cluster head node consumes energy in receiving the data from other member nodes, then aggregating these collected data and in transmitting the aggregated data. Member nodes send the data packets of length l. So energy consumption of a cluster head in a round is:

$$\begin{aligned} E_{CH/Round} &= X_{frames/round} \times E_{CH/frame} \\ &= X_{frames/Round} \times (\frac{N}{k} l E_{elec} + \frac{N}{k} l E_{BF} + l\epsilon_{two-ray} d^4_{toBS}) \end{aligned} \tag{4.2}$$

Where, E_{elec} is the energy spent to receive a packet, E_{BF} is the energy spent in performing data aggregation and d_{toBS} is the distance of cluster head node to base station. As the base station is situated far outside the field, energy consumed in transmission of aggregated data to base station depends on $d^4{}_{toBS}$ (i.e. Two-ray ground reflection propagation model is applied).

Member nodes only consume energy to communicate with cluster head. Hence energy consumption of a non-cluster head node in a round is:

$$\begin{aligned} E_{Non-CH/Round} &= X_{frames/Round} \times E_{Non-CH/frame} \\ &= X_{frames/Round} \times (l E_{elec} + l\epsilon_{friss} \frac{1}{2\pi} \frac{M^2}{k}) \end{aligned} \tag{4.3}$$

A node should be active for performing the role of cluster head in one round and member node in other rounds. In N/k rounds each node is selected as cluster head once in every round. To be a cluster head a node should have enough energy for at least one round and non-cluster head for (N/k - 1) rounds. Therefore

$$E_{CH/Round} + (\frac{N}{k} - 1) E_{Non-CH/Round} = E_{Start} \tag{4.4}$$

So, from equation (4.2), (4.3) and (4.4), we have

$$X_{frames/round} = \frac{E_{Start} \backslash l}{(\frac{N}{k} E_{elec} + \frac{N}{k} E_{BF} + \epsilon_{two-ray} d^4_{toBS}) + (\frac{N}{k} - 1)(E_{elec} + \epsilon_{friss} \frac{1}{2\pi} \frac{M^2}{k})} \tag{4.5}$$

Let, each cluster has N/k nodes, and l-bit of data takes t_{msg} = l/Rb seconds over a bandwidth of Rb bits/sec. Then, the total frame length is:

$$t_{frame} = \frac{N}{k}\frac{l}{Rb} \tag{4.6}$$

From Equation (4.1), (4.5) and (4.6), we have the length of round-time:

$$t_{round} = \frac{1}{Rb}\frac{N}{k}\frac{E_{Start}}{\left[\left(E_{CH/frame}\right) + \left(\frac{N}{k} - 1\right)\left(E_{Non-CH/frame}\right)\right]} \tag{4.7}$$

But due to network dynamics, nodes die randomly and at different time. Let there are N_{Alive} nodes at the end of (r-1)th round. Thus for rth round, there are N_{Alive}/k nodes per cluster. Thus, the total frame time length for rth round is:

$$t_{frame} = \frac{N_{Alive}}{k}\frac{l}{Rb} \tag{4.8}$$

So, equation (4.7) should be as

$$t_{round} = \frac{1}{Rb}\frac{N_{Alive}}{k} \times \frac{E_{Start}}{\left[\left(E_{CH/frame}\right) + \left(\frac{N}{k} - 1\right)\left(E_{Non-CH/frame}\right)\right]} \tag{4.9}$$

Solution for constant round-time problem is to calculate round-time dynamically using the proposed equation 4.9, i.e. round-time depends on the number of alive nodes[1] in the network. As the network progresses, nodes start dying in the field and hence the proposed *Dynamic Round-Time* solution has shorter round-time than the initial length of round-time. The analysis of proposed solution of dynamic round-time exposed that the round-time has been very short when the network has few alive nodes left in the last rounds. It decreases the performance of proposed solution because a shorter round-time causes frequent re-clustering that exhausts the remaining energy of nodes very quickly. So a further enhancement is also required for the proposed dynamic round-time solution that takes care of last few nodes of the network so that the performance of network can be improved significantly as whole.

[1]All nodes have the knowledge about the number of alive nodes in the network with application of any existing technique - WBCHN [90] or some centralized manner [58]. Energy consumption of nodes by these exchange messages are not considered in our work.

With the performance analysis of LEACH and its variant clustering algorithms, it can be concluded that at the time of 50% node death 70% - 80% of total energy of network has been consumed (Alive nodes of last rounds have little remaining energy). So, a further enhancement over the proposed dynamic round-time solution has been suggested names as *Network Adaptive Round-time.* The proposed network adaptive round-time solution has dynamic round-time according to equation 4.9 up to the death of 50% nodes and has constant round-time after that. As described earlier, at the time of 50% node death most of the energy of the network has been consumed so we fixed the round-time after that benchmark performance metric. Proposed network adaptive round-time solution figured length of round-time according to number of alive nodes present in the network so at the death of 50% nodes the length of round-time is exact half of the initial length. Therefor, in the proposed network adaptive round-time solution, the length of round-time is half of initial value when it has been fixed for last few rounds after 50% remaining nodes. Subsequently, the proposed network adaptive round-time solution integrates round-time to network dynamic and also takes care of trade-off of longer and shorter round-time for the last few remaining nodes.

Block diagram of proposed solution of network adaptive round-time is shown in figure 4.1. For the proposed network adaptive round-time solution, number of alive nodes in the network is known to each node. Therefor, in the round-time calculation phase length of round-time is figured out.In the proposed network adaptive round-time solution, round-time is not changed after 50% node death consequently round-time is half of initial value. In the clustering phase, cluster head selection, cluster formation, TDMA scheduling and data processing is done as in LEACH [51].

4.3 Performance Evaluation of Proposed Solution for Round-Time

The performance of network adaptive round-time solution is compared with LEACH [2]. We also inspect the dynamic round-time solution that calculates round-time dynamically till the death of last node. We review the *network lifetime* and *data received at base*

[2]For the simulation results, dead nodes are figured out by checking the energy level and if energy level of node is less than some predefined level, nodes are considered as dead. Eventually, we have the number of dead nodes in the network, and we have the information about number of alive nodes in the network

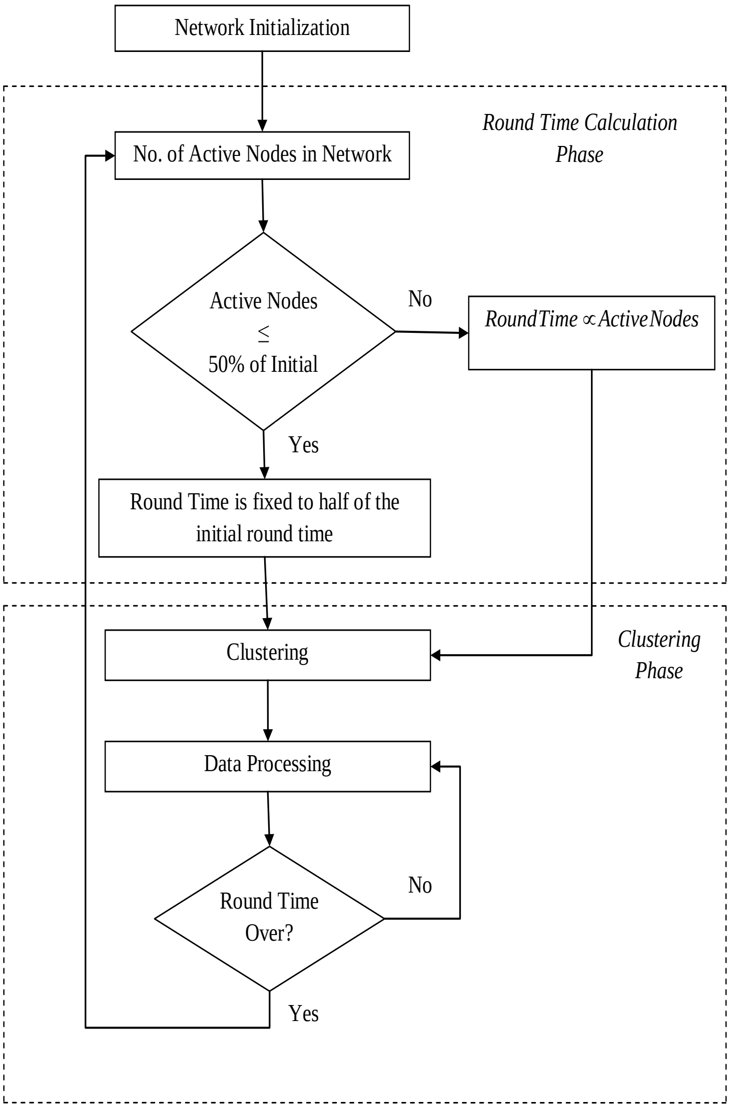


FIGURE 4.1: Block Diagram of Proposed Solution

station metrics. The proposed solution change the round-time after the death of the first node so simulation time of 50% node death and 95% node death are considered for network lifetime.

Three network topologies each with different area and number of nodes, 50×50 m^2 with 50 nodes, 100×100 m^2 with 100 nodes and 150×150 m^2 with 200 nodes, are generated and simulated in NS-2. Other simulation parameters are listed in table 4.1. Radio parameters used for simulation are described in chapter 3. Initial round-time here for network topologies, 50×50 m^2 with 50 nodes, 100×100 m^2 with 100 nodes and 150×150 m^2 with 200 nodes, is 10 sec, 20 sec and 40 sec. respectively. In LEACH, round-time is

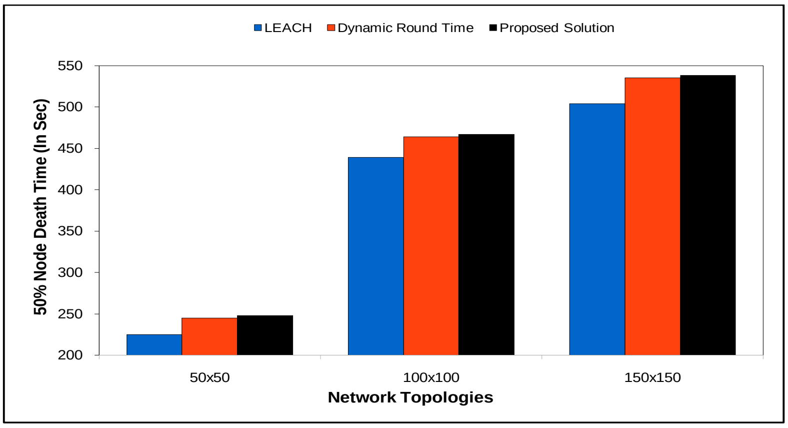


FIGURE 4.2: Comparison of 50% Node Death

fixed while in proposed solution, it varies in accordance with the number of alive nodes.

TABLE 4.1: Simulation Parameters and Values

Parameters	**Values**
Network Area	50×50 m^2, 100×100 m^2, 150×150 m^2
Number of Nodes	50, 100, 200
Header Packet Size	25 Bytes
Data Packet Size	500 Bytes
Initial Energy	2 Joules

Figure 4.2 shows 50% node death time for LEACH, dynamic round-time solution and network adaptive round-time solution. In all three network topologies, there is significant improvement for network adaptive round-time solution and dynamic round-time solution over LEACH. There is improvement of 13%, 7% and 10% for 50×50 m^2 with 50 nodes, 100×100 m^2 with 100 nodes and 150×150 m^2 with 200 nodes topologies in the case of proposed network adaptive round-time solution over LEACH respectively. Both network adaptive round-time solution and dynamic round-time solution have the same duration for round-time till 50% node death; hence have almost the same improvement (there are small variations due to the random nature of NS-2 simulator).

Figure 4.3 shows network lifetime (95% node death) for LEACH, dynamic round-time and network adaptive round-time solution. There is an improvement of 7% for 50×50 m^2 topology, 6.5% for 100×100 m^2 and 150×150 m^2 topologies for the network adaptive round-time solution. Dynamic round-time solution does not have much improvement while having improvement for 50% node death because the round-time duration becomes very short and the network has the extra burden of frequent re-clustering. While LEACH has constant round-time for the nodes having very less remaining energy which is long enough to complete for these nodes. This also affects the data units received at base station.

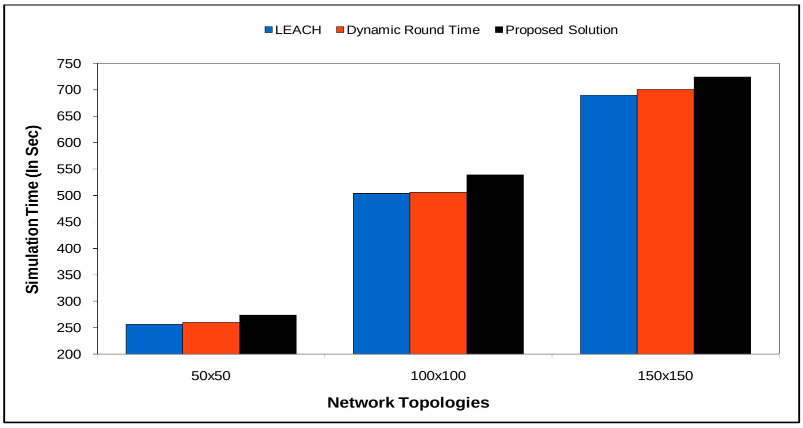


FIGURE 4.3: Comparison of Network Lifetime

Total successfully received data units at the base station are shown in figure 4.4. The network adaptive round-time solution has improvement over LEACH and dynamic round-time solution. The network adaptive round-time solution extends the network lifetime so the network has more time to sense the area and then send the data to base station. There is increase of 7%, 5% and 6% for 50×50 m^2 with 50 nodes, 100×100 m^2 with 100 nodes and 150×150 m^2 with 200 nodes topologies in case of network adaptive round-time solution over LEACH respectively. In LEACH, round-time is long enough to complete for less remaining energy nodes of the progressed network, so the data of that cluster is not received at base station. In dynamic round-time solution, progressed network has the burden of frequent re-clustering.

4.4 Proposed Solution for Cluster Head Selection

This section anatomizes the effect of cluster head position in the cluster on intra-cluster communication distance. Cluster head selection for less intra-cluster communication distance is illustrated further in the section.

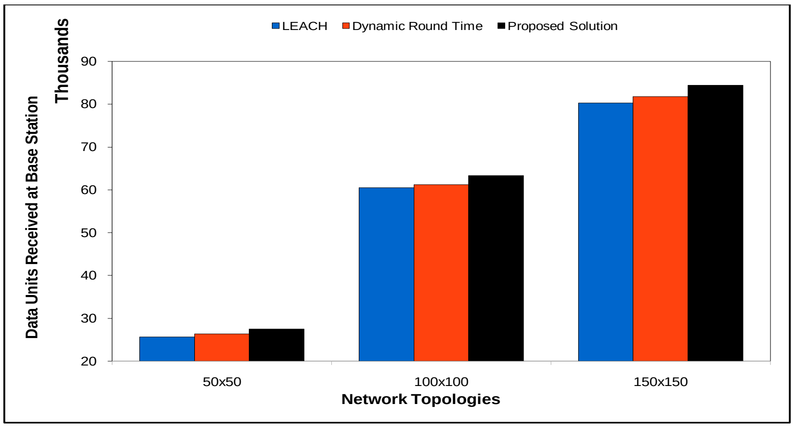


FIGURE 4.4: Data Units Received at Base Station

4.4.1 Intra-cluster Communication Distance

Clustering algorithms groups the nodes in clusters. A cluster head first receives data from cluster head and then transmit this aggregated data to base station. The energy consumed by the communication of the data depends on the distance between two nodes. Communication in clustering algorithms can be classified in two categories: *Intra-cluster communication* and *Inter-cluster communication*. Distances involved in communication can also be categorised as: *Intra-cluster communication distance* and *Inter-cluster communication distance*. Therefore, intra-cluster communication distance of a cluster can be defined as the sum of the distance of all member nodes of a cluster to cluster head. So,

$$Intra - cluster\ communication\ distance = \sum_{i=1}^{N} Dist(i, CH) \tag{4.10}$$

where N = number of member nodes in cluster and,

Dist(i,CH) = function to describe distance of node i to cluster head.

Energy efficiency of a cluster depends on the intra-cluster communication distance. Data transmission phase is made longer than cluster set-up phase to avoid the burden of frequent re-clustering. Data transmission phase has both inter and intra-cluster communication repetitively that makes intra-cluster communication distance much higher than inter-cluster communication distance. So intra-cluster communication consumes most of the energy of network.

4.4.2 Effect of Cluster Head Selection on Intra-cluster Communication Distance

Intra-cluster communication distance is the measure of efficiency of a cluster. It depends upon the position of cluster head in the cluster. We posit a wireless sensor network of 50 nodes over an area of 50×50 m^2 area as depicted in figure 4.5.

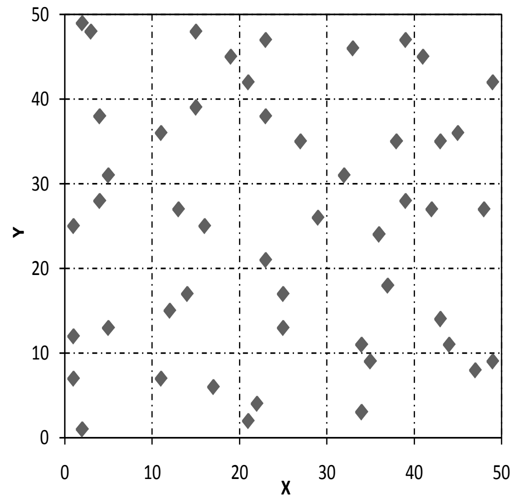


FIGURE 4.5: Wireless sensor networks of 50 nodes

Let there is only one cluster head in the network, i.e. there is only one cluster. Intra-cluster communication distance is calculated for each node considering it as cluster head and other as member nodes. We show intra-cluster communication distance of few clusters which have the highest and the lowest distance in table 4.2.

TABLE 4.2: Intra-cluster communication distance of clusters

CH with maximum intra-cluster communication Distance		**CH with minimum intra-cluster communication Distance**	
Node Location(x,y)	**Total Distance**	**Node Location(x,y)**	**Total Distance**
1,49	976	15,15	491
47,39	938	25,15	499
46,7	877	12,22	526
46,6	865	30,25	523
2,1	855	32,31	547

For the deployed network, table 4.2 reveals that cluster head positioned at the border of cluster has more intra-cluster communication distance while cluster head positioned near to the center of cluster has less intra-cluster communication distance. So the position of cluster head node in cluster plays an important role for the energy efficiency aspect of clustering algorithms. Cluster head selection should be amended by clustering algorithms to have clusters with less intra-cluster communication distance.

4.4.3 Proposed Solution

Various approaches- genetic algorithm, particle swarm optimization, etc., are adopted by clustering algorithms to redress cluster head selection and cluster formation. But most of these schemes cannot be implemented distributed because these schemes necessitate complex and intelligent computing. Sensor nodes are resource constraints and do not posses circuitry to perform these complex computing. Computing also consumes energy of a node. A new cluster head selection scheme must be incorporated that does not require complex computing and also has energy efficient clusters. Cluster head selection approach presented in this chapter can be implemented distributed and has efficient clusters without doing complex computing.

Sensor nodes are deployed randomly in the field, i.e. the positions of sensor nodes are not pre-engineered for deployment. Nodes can be equipped with GPS (Global Positioning System) to get the information about the position but it increases the cost of sensor nodes. Localization algorithms [110, 111] can be applied to have the information about the location of nodes. Since it will be applied once because sensor nodes are not moving in the field, it will not affect energy efficiency feature of clustering algorithms. In the proposed solution, nodes own information about the location by means of localization algorithm. The proposed solution divides the area of the network into two parts: *Border Area* and *Inner Area* as shown in figure 4.6.

Let d be the partitioning distance to divide the network area. The area starting from the boundary of the field and up to the distance d is named border area and the remaining area is known as inner area. In the proposed solution, border area nodes cannot participate in cluster head selection procedure. Only inner area nodes can participate for the cluster head role. Border area nodes will be always member nodes in each round. In each round, new clusters are configured therefore clusters do not have fix boundary but the boundary of the network area is fix. As discussed in section 4.4.2, clusters having

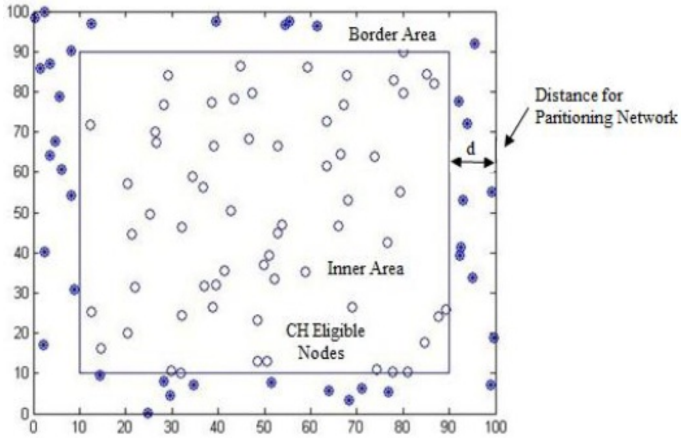


Figure 4.6: Division of Network Area

cluster head node positioned away from the center have more intra-cluster communication distance consequently clusters formed around the node of border area have more intra-cluster communication distance. In the proposed solution, cluster heads are always selected from the inner area so cluster heads are always near to the center of cluster.

Cluster head node dissipates much higher energy than a member node. Only inner area nodes compete for the role of cluster head and hence exhaust more energy than border area nodes but the proposed solution still has a well load balanced network. Inner area nodes as member nodes communicate for short distance because cluster head is always from the inner area, so consume less energy. Border area nodes communicate over a long distance but are never in the role of cluster head. So the proposed solution conserves the load balance in the network.

Partitioning distance (d) is pivotal point for the proposed solution. Inner area will be very small for higher value of d, therefore there will be less number of nodes eligible for becoming cluster head. These nodes will dissipate their energy very quickly as cluster head role is more energy consuming. There will be no alive node for the role of cluster head therefore, the network will sustain for short time. While a lower value of d does not have enough change for clustering algorithms. In a network of 100 sensor nodes deployed uniformly over 100×100 m^2 area, the number of border area nodes and inner area nodes are shown in table 4.3 for various values of d. The effect of the value of d on the performance of proposed clustering algorithm solution is analyzed in section 4.5.1.

Table 4.3: Number of nodes in Border Area and Inner Area for Different Values of d

	d=0	d=5m	d=10m	d=15m	d=20m	d=25m	d=30m
Border Area Node	0	12	36	48	64	72	84
Inner Area Node	100	88	64	52	36	28	16

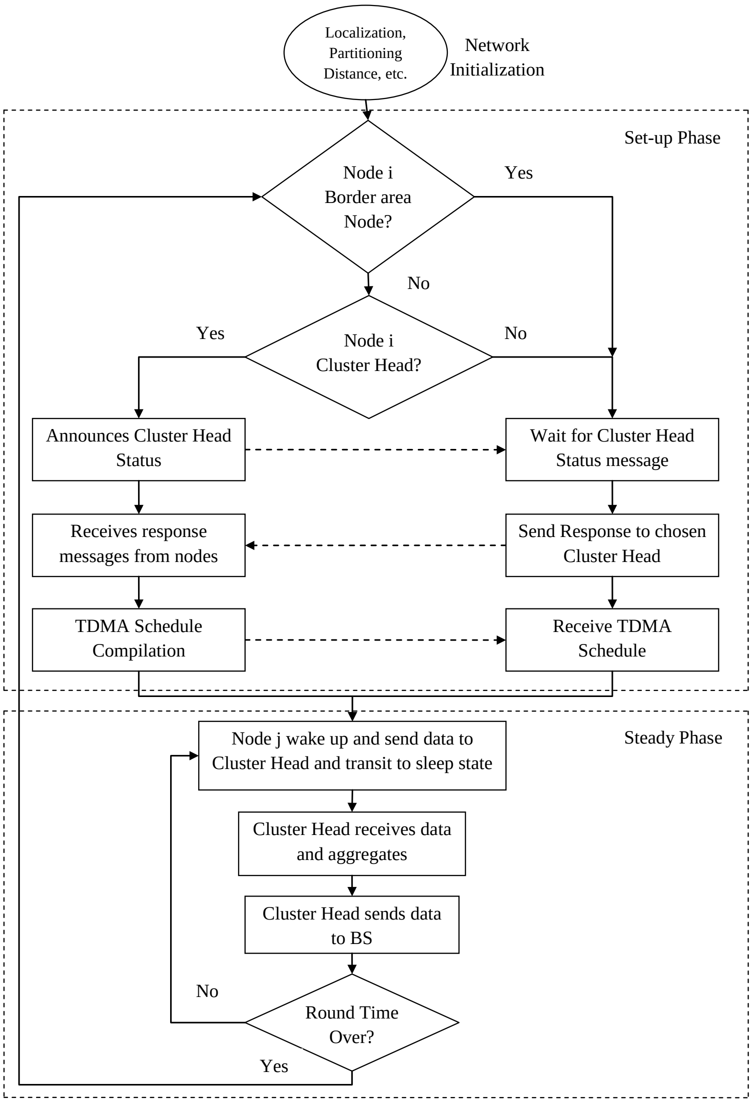


Figure 4.7: Block Diagram of Proposed Solution

Block diagram of the proposed solution for cluster head selection for proposed clustering algorithm is demonstrated in figure 4.7. Network start-up, localization of nodes and partitioning distance is done in network initialization. Network start-up deals with constructing the network by finding and connecting nodes together, assigning unique ID to each sensor nodes, etc. Localization algorithm to find the location of sensor nodes in the deployed area is applied once because nodes are not mobile in the field. Knowledge of partitioning distance is imposed to all sensor nodes. Localization and partitioning distance recognize nodes for border area and inner area. In set-up phase, only inner area nodes get involve for cluster head selection while border area nodes are always member nodes. Cluster head nodes are selected based on a probabilistic approach. A random number between 0 and 1 is selected by node. A threshold (Th) is calculated for each round according to the equation (4.11). If this random number for a node is less than threshold, that node is selected as cluster head.

$$T(n) = \begin{cases} \frac{p}{N_{INNER} - p \times (r mod \frac{N_{INNER}}{p})} & \text{if } n \in G \\ 0 & \text{otherwise} \end{cases} \tag{4.11}$$

where r = number of current round,

p = number of optimal cluster heads, and

N_{Inner} = number of inner area nodes.

Selected cluster heads announce their cluster head status to the network and wait for the cluster joining request from other nodes. Nodes in the network receive advertisement message from cluster heads and send join-request to the nearest cluster head by comparing receive signal strength (RSS) of messages. Cluster heads receive join-request and after that compile and broadcast TDMA schedule of respective clusters. In steady phase, node sends data to cluster head in assigned time slot in TDMA schedule. Rest of the time, node is in sleep state to conserve energy. At the end of the frame, cluster head aggregates the collected data and transmit the aggregated data to base station. Round-time calculated according to section 4.2 is implemented.

4.5 Performance Evaluation

To anatomize the performance of proposed clustering solution with LEACH, both approaches are implemented in NS-2. Following metrics are considered: *Node Death Rate, Energy Consumption Rate, Network Lifetime (First Node Death, 50% Node Death and*

95% Node Death) and *Data Units Received at Base Station*. Simulation parameters and values listed in table 4.1 are considered.

4.5.1 Distance for Partitioning the Application Area

Partitioning of network area - border area and inner area, is cardinal issue for the performance of the proposed solution. As explained in section 4.4.3, partitioning distance (d) determines the number of nodes in border area and inner area. Varied values of d are examined and the performance of the proposed solution is evaluated for a network of 100×100 m^2 with 100 uniformly distributed nodes. Node death rate for varying values of d is depicted in figure 4.8. Less value of d does not change the network scenario so has very little effect as in the case of d=5m. High value of d makes less number of nodes in inner area so the network has very short epoch for inner area nodes therefore nodes consume their energy very quickly. It is the case of 20m, 25m and 30m. It can be analyzed from figure 4.8, d=10m has the lowest node death rate. Therefore, for the rest of the work value of d is set 10m.

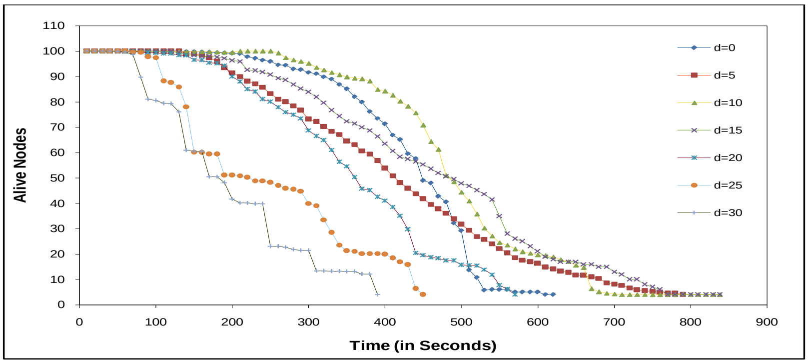
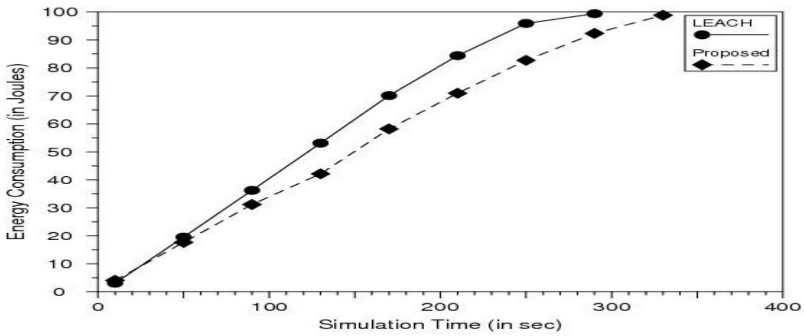


FIGURE 4.8: Node Death Rate for 100×100 m^2 with 100 Nodes Network for different value of d

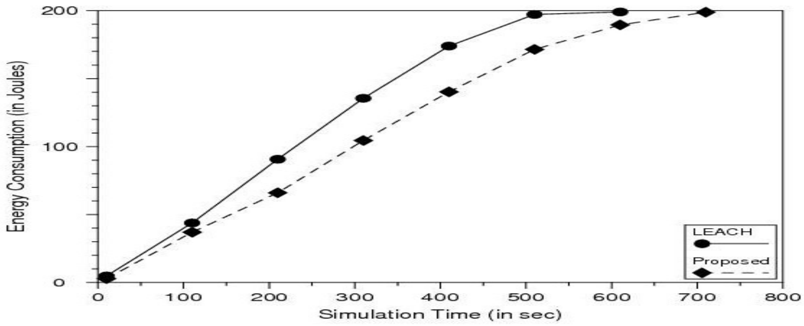
4.5.2 Simulation Results

a. Energy Consumption Rate

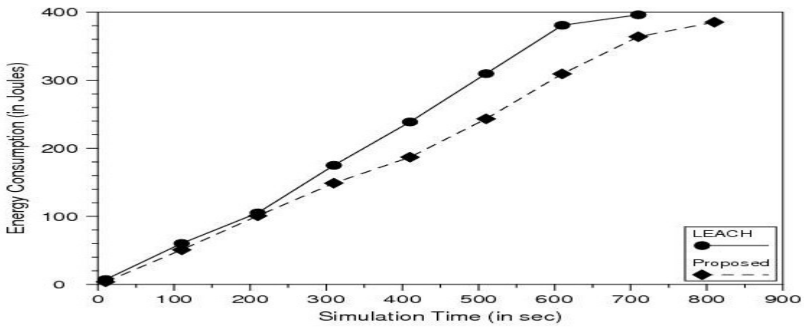
Figure 4.9(a-c) shows energy consumption rate for all three considered network topologies. It is analyzed that in all three network topologies, rate of energy consumption for the proposed solution is lower than that in LEACH. The proposed scheme has less intra-cluster communication distance as compared to LEACH, so clusters in the proposed solution have energy efficient communication.



(a) Energy consumption for 50×50 m^2 area with 50 nodes



(b) Energy consumption for 100×100 m^2 area with 100 nodes

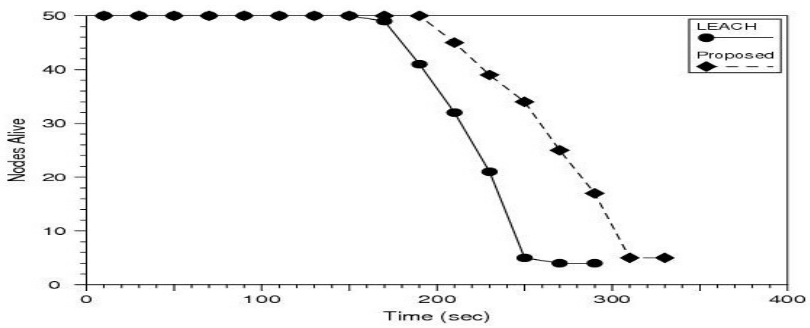


(c) Energy consumption for 150×150 m^2 area with 200 nodes

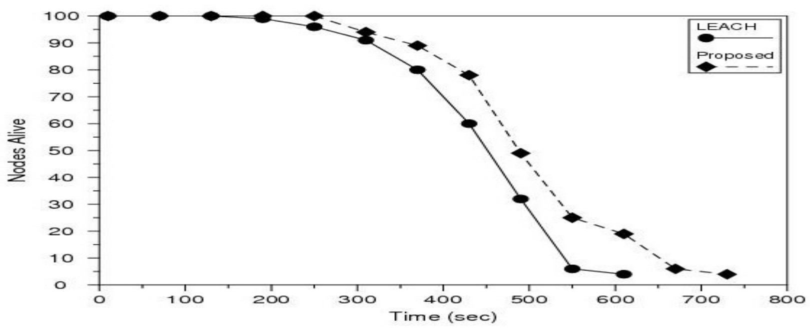
FIGURE 4.9: Energy consumption over Time

b. Node Death Rate

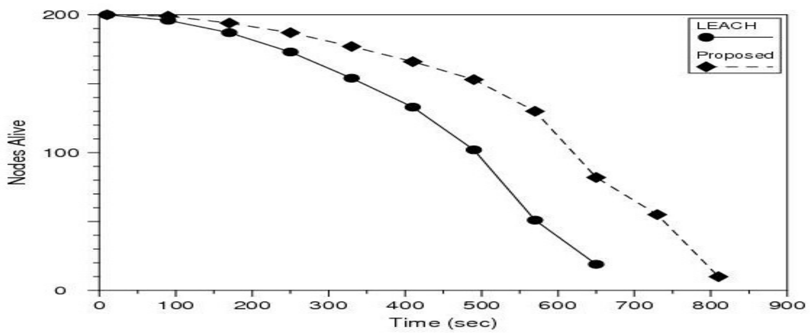
As examined from figure 4.9(a-c), the proposed solution has lower energy consumption rate as compared to LEACH consequently the proposed solution should have lower node death rate as compared to LEACH. Figure 4.10(a-c) shows node death rate of all three network topologies. The number of alive nodes over the simulation time in the proposed solution are always more than of LEACH as expected. Proposed solution has longer



(a) Nodes Alive vs. Simulation Time for 50×50 m^2 area with 50 nodes



(b) Nodes Alive vs. Simulation Time for 100×100 m^2 area with 100 nodes



(c) Nodes Alive vs. Simulation Time for 150×150 m^2 area with 200 nodes

Figure 4.10: Number of Nodes Alive over Time

stable region as compared to LEACH that means the network is better load balanced for proposed solution compared to LEACH.

The proposed solution extends stable region by 25% and 32% for 50×50 m^2 with 50 nodes, 100×100 m^2 with 100 nodes topologies respectively. The proposed solution places the cluster head near to the center of cluster therefore, nodes consume energy more efficiently. The proposed solution also prolongs unstable region of the curve. In case of 150×150 m^2 with 200 nodes topology, stable region is short as compared to other topologies because performance of clustering algorithm softens as the network grows both in area and number of nodes. Energy efficient clustering algorithm for large network is required but is out of scope for the work of this thesis.

c. Network Lifetime

As analyzed from figure 4.9(a-c) and 4.10(a-c), the proposed solution has lower energy consumption and node death rate as compared to LEACH. The proposed solution consumes energy of nodes efficiently. So, the proposed solution should have better network lifetime as compared to LEACH. Figure 4.11, 4.12 and 4.13 presents comparison of network lifetime of the proposed solution and LEACH.

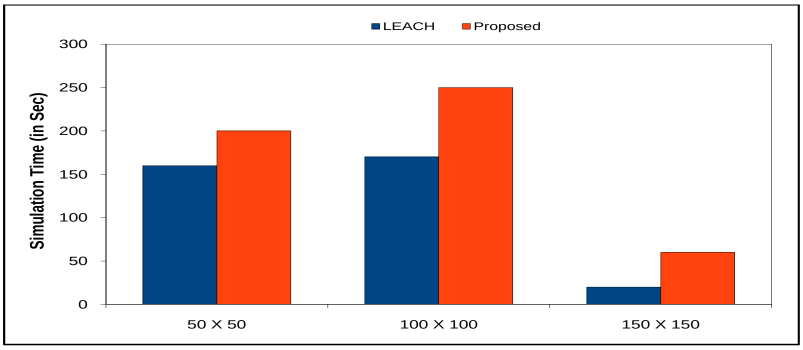


Figure 4.11: First Node Death vs. Simulation Time

Figure 4.11 demonstrate the time of first node death for all three network topologies. As analyzed in figure 4.10(a-c), proposed solution has longer stable region that means the time of first node death is longer for proposed solution as compared to LEACH. There is an increase of 25% and 32% for the first node death time for 50×50 m^2 with 50 nodes, 100×100 m^2 with 100 nodes topologies respectively. There is also significant improvement for the topology 150×150 m^2 with 200 nodes.

Figure 4.12 shows time of 50% node death for all three network topologies. Node death rate of the proposed solution is lower than LEACH so that the proposed scheme provides time of 50% node death which is longer than that of LEACH. There is improvement of 20%, 10% and 20% in 50% node death time for 50×50 m^2 with 50 nodes, 100×100 m^2 with 100 nodes and 150×150 m^2 with 200 nodes topologies in case of proposed solution over LEACH respectively.

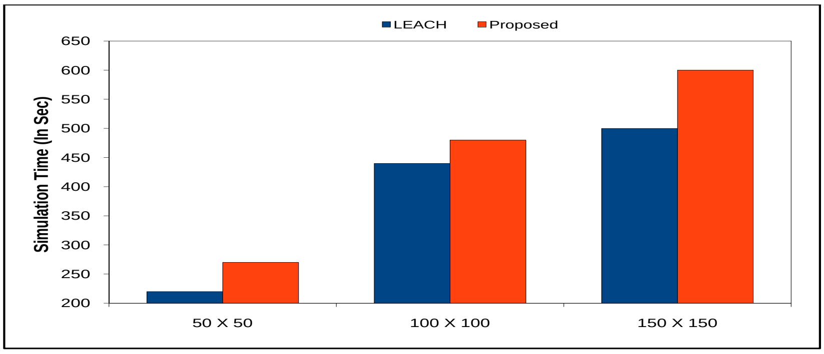


FIGURE 4.12: 50% Node Death vs. Simulation Time

The time of 95% node death is shown in figure 4.13 for all three network topologies. As examined from figure 4.9(a-c), unstable region of the proposed solution is also longer than LEACH, consequently the network in the proposed solution sustains for a longer time. Hence there is significant increase in 95% node death time for the proposed solution over LEACH. Therefore, the proposed solution performs much better than LEACH to prolong the network lifetime. The network is better load balanced in proposed solution compared to LEACH.

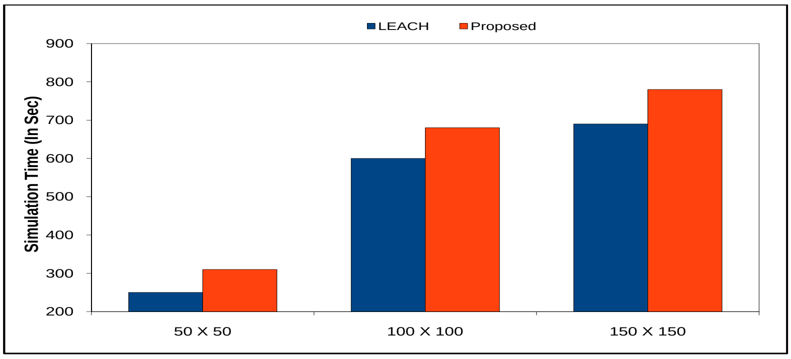


FIGURE 4.13: 95% Node Death vs. Simulation Time

d. Data Gathering

The proposed solution has lower node death and energy consumption rate over LEACH so has a prolonged network lifetime. Increased network lifetime means network has more time at its disposal to sense the area and to transmit the data to base station. Figure 4.14 depicts the number of data units successfully received at base station during the complete network lifetime for all three considered network topologies. As can be analyzed, more data units are received at base station with proposed solution as compared to LEACH. There is increase of 14%, 8% and 12% for 50% node death time for 50×50 m^2 with 50 nodes, 100×100 m^2 with 100 nodes and 150×150 m^2 with 200 nodes topologies for proposed solution over LEACH.

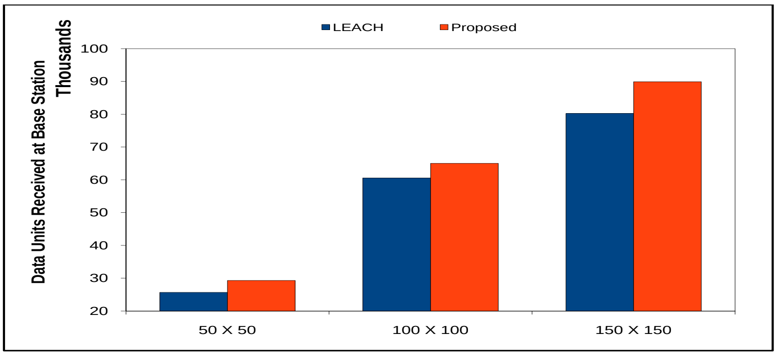


Figure 4.14: Data Units Received at Base Station vs. Simulation Time

So it can be summarized that the proposed solution has better performance when compared with LEACH in terms of node death rate, energy consumption rate, network lifetime and data units successfully received at base station.

4.6 Conclusion

A clustering approach is presented in this chapter that has network adaptive round-time and energy efficient cluster head selection solution. Presented solution figures the round-time based on the number of alive nodes in the field. After the death of 50% nodes, round-time is made constant because nodes have very less remaining energy and a further short round-time will dissipate more energy due to the added task of re-clustering.

Clusters with cluster head positioned near to the center of the cluster have less communication distance hence are appraised energy efficient. The proposed cluster head selection scheme places cluster head near the center of the cluster. The network area has two parts: border area and inner area. Only inner area nodes can contend for the role of cluster head. Border area nodes are always member nodes. Partitioning distance to divide the network area is vital for the performance of the proposed solution of cluster head selection. It is obvious from the Simulation results that the proposed solution outperform the traditional clustering algorithm LEACH.

Chapter 5

Exploiting Heterogeneity of Sensor Nodes to Protract Wireless Sensor Network Lifetime

Heterogeneity of sensor nodes gives a facelift to the performance of wireless sensor networks. Work of this chapter presents a clustering approach that makes the most of heterogeneity of sensor nodes which is either due to dynamic nature of network or a percentage of the total nodes have extra energy. Proposed solution takes into account both remaining node energy and average network energy at the beginning of current round for cluster head selection.

5.1 Introduction

Sensor networks can be either homogenous or heterogeneous. A homogenous networks is called so because all the nodes are equipped with same circuitry (i.e. nodes have equal amount of energy, have equal communication and processing capabilities). In heterogeneous networks, nodes do not have same circuitry (i.e. nodes have different communication and processing capabilities and also have different amount of energy).

It is assumed by the traditional clustering algorithms for a homogeneous network (same amount of energy for each node) that homogeneous nature of nodes is retained throughout the network lifetime [51, 52]. But it becomes cumbersome to maintain the homogeneity of the nodes as the network proceeds forward towards the task completion. The amount of energy dissipated by any node in the network depends on the job assigned to the respective node. As the energy consumed by a cluster head node is higher than that consumed by the member nodes. In most of the algorithms, cluster heads are supposed to be in wake-up state all the time so as to receive the data anytime while on the other hand a member node remains in sleep state and come into wake-up state only in its assigned time slot thereby using less amount of energy. The only node which communicates with the base station, located at a long distance, is the cluster head. Therefore a cluster head spends energy in larger amount when compared to member nodes. In any network member nodes are located from a cluster head at different distances and hence each member operates at different energy levels at any moment of time.

With the help of previous simulation results, we explain how an initially homogeneous network transmutes to heterogeneous. A homogeneous network with 50 nodes with 4 cluster heads is considered. Energy consumption of various cluster heads and their respective member nodes is as shown in Table 5.1. It is evident from the table that a cluster head consumes more energy than its member counterparts. Cluster heads do not dissipate same amount of energy for they have different number of member nodes and their distance from the base station also varies. Energy spent by a member node depends on its distance from the cluster head, larger the distance more are the energy requirements. Hence an initially homogeneous network starts behaving as a heterogeneous one as the operation proceeds.

Energy Consumption of Cluster Head Node (In Joules)	**Energy Consumption of Member Nodes (In Joules)**
0.354	0.012,0.011,0.036
0.382	0.015,0.021,0.023
0.214	0.021,0.020,0.036
0.386	0.019,0.027,0.045

Table 5.1: Energy Consumption of cluster head and member nodes

Nowadays, heterogeneous wireless networks have achieved a great importance as heterogeneity of nodes facilitates the network performance without demanding additional costs.

Wireless sensor networks have three types of hardware heterogeneity (ref. M.Yarvis et al [26]) *Computational*, *Link* and *Energy*. Computational heterogeneous nodes are equipped with more powerful computational circuitry and added storage units. Complex data computations are performed at these nodes. Link heterogeneous nodes enjoy high quality communication link with base station and hereby provide better link. Latency in the network is reduced by means of these heterogeneous nodes and they also provide base station with highly reliable data. High power batteries are employed in the energy heterogeneous sensor nodes. M.Yarvis et al [26] also states that performance of the network is significantly improved when these heterogeneous nodes are employed. Among these heterogeneities Energy heterogeneity has considerable importance as computational and link heterogeneities can deteriorate the network lifetime when used without energy heterogeneity because this might cause energy of the nodes to dissipate rather quickly. Therefore it is of paramount importance to consider the sensor-node-heterogeneity caused either by network dynamics or due to the initial set-up.

Consideration of heterogeneity of sensor nodes either caused by dynamics of network or due to the initial set-up of network is compelling for energy efficient clustering.

5.2 Proposed Solution

5.2.1 Problem Formulation

A network is load balanced by rotating the role of cluster head among the nodes. p_{opt}, the optimal probability of cluster heads per rounds is defined by the spatial density of the nodes in the network [13]. A cluster node act as a cluster head once in $1/p_{opt}$ rounds. These numbers of rounds is termed as epoch of the nodes as cluster head.

In LEACH, a nodes plays the role of a cluster head only once in an epoch. A number between [0,1] is randomly selected by each node and a node is selected as cluster head if this selected number is less than the threshold. This threshold value depends on the current round number which is calculated as:

$$T(n) = \begin{cases} \frac{p_{opt}}{1-p_{opt}\times(r mod \frac{1}{p_{opt}})} & \text{if } n \in G \\ 0 & \text{otherwise} \end{cases} \tag{5.1}$$

where r is the current round number, p_{opt} is optimal cluster head probability and G is the set of nodes that are not selected as cluster head in that epoch. In LEACH, all nodes have equal probability of being selected as cluster head at the beginning of round. LEACH does not ponder energy heterogeneity of nodes for cluster head selection.

where r = current round number

p_{opt} = optimal cluster head probability

G = nodes not selected as head in an epoch

The probability of a node being selected as a cluster head at the beginning of each round is equal in LEACH. And LEACH does not take into account the energy heterogeneity of the nodes.

In SEP protocol different probabilities for cluster head selection are employed to handle the heterogeneity of the nodes - *advance nodes* and *normal nodes*. But it, too, does not consider the residual energy of nodes while selecting the cluster heads. It might result in a low-energy node being selected as a cluster head when there are high-energy nodes are available in the cluster. Residual energies of different nodes for a round r, random number and threshold value selected by any node to take part in the selection process for LEACH and SEP are as illustrated in the table 5.2.

With LEACH algorithm B, D and F are chosen as cluster head for the round r. And as shown in the table nodes D and F are selected as heads though there are high residual energy nodes available. Whereas with SEP B, C and F are chosen as cluster head for the said round r and in this case C and F have low residual energy. LEACH and SEP select a cluster head depending on the random number selected by the node and do not consider the remaining energy of the respective node. The given solution proposes to exploit the heterogeneity of nodes for efficient cluster head selection.

TABLE 5.2: Cluster Head Selection in LEACH and SEP

Node	**Residual Energy (Joule)**	**LEACH**		**SEP**	
		Thresh -old	**Random Number**	**Thresh -old**	**Random Number**
A*	2.0	0.5	0.72	0.45	0.62
B*	1.5	0.5	0.51	0.45	0.41
C	1.0	0.5	0.64	0.6	0.55
D	0.5	0.5	0.43	0.6	0.72
E	1.2	0.5	0.65	0.6	0.88
F	0.4	0.5	0.41	0.6	0.45
G*	1.7	0.5	0.83	0.45	0.87

*(Advance nodes in case of SEP, node heterogeneity by network dynamics in LEACH)

5.2.2 Proposed Solution

Let N, initially, homogeneous are uniformly distributed over an M×M region. In the proposed scheme, nodes select a random number which is not a pure random number as done in conventional clustering algorithms. A random number is derived from a pure random number after considering residual energy of nodes ($E_{residual}$) and average network energy ($E_{average}$) for that round. A random value between 0 and 1 is selected by the node and this value is further processed to a new value as:

$$value(i) = random - value \times (1 - \frac{E_{residual}(i) - E_{average}}{E_{average}}) \tag{5.2}$$

According to equation (5.2) nodes with high residual energy than average network energy have lower random value and vice versa. Therefore a high residual energy node has higher probability of being selected as a cluster head than lower residual energy nodes in that round. Equation (5.1) is used to calculate the threshold value. Table 5.3 shows residual energies of different nodes for a round r, random number and threshold value selected by any node to take part in the selection process for LEACH and proposed scheme for a network which is initially homogeneous.

It is evident from the table (5.3) that proposed solution selects high residual energy nodes as cluster heads for round r. Nodes A,B and G with higher residual energy have low new random number value. Hence nodes A, B and G are selected as cluster heads.

Each node in cluster head selection process takes average energy of network as reference energy. So

$$E_{average} = \frac{1}{N}\sum_{i=1}^{N} E_{residual}(i) \tag{5.3}$$

Total approximate number of rounds (R) for network assuming that all the nodes are uniformly deployed can be evaluated as:

$$R = \frac{E_{Total}}{E_{Round}} \tag{5.4}$$

A cluster head node spends its energy in the following tasks: receiving data from other member nodes, aggregating that data and transmitting it to the base station. L bit data is transmitted from member node to cluster head and it is assumed that there is no line of sight path available from cluster head to base station. E_{Round} for k clusters is given by [20]:

$$\begin{aligned} E_{Round} &= X_{frame/round} \times L(2NE_{elec} + NE_{DA} + k\epsilon_{two-ray}d^4_{toBS} \\ &\quad + k\epsilon_{frss-amp}d^2_{toCH}) \end{aligned} \tag{5.5}$$

Therefore average energy of the network for round r can be approximated as (from equations 5.3, 5.4 and 5.5):

$$E_{average} = \frac{1}{N} \times E_{Total} \times (1 - \frac{r}{R}) \tag{5.6}$$

We have taken double the value of R to deal effectively with the network dynamics.

TABLE 5.3: Cluster Head Selection in LEACH and Proposed solution for homogeneous network

Node	**Residual Energy (Joule)**	**LEACH**		**Proposed Solution**	
		Thresh -old	**Random Number**	**Thresh -old**	**Random Number**
A	2.0	0.5	0.72	0.5	0.44
B	1.5	0.5	0.51	0.5	0.36
C	1.0	0.5	0.64	0.5	0.69
D	0.5	0.5	0.43	0.5	0.72
E	1.2	0.5	0.65	0.5	0.61
F	0.4	0.5	0.41	0.5	0.75
G	1.7	0.5	0.83	0.5	0.49

5.2.3 Heterogenous Network

Advance nodes and normal nodes are two different types of nodes. Let m is the fraction of advance nodes which have α times more energy. So there are (1-m.N) normal nodes with initial energy E_{Start}. Therefore total network energy is given by:

$$\begin{aligned} E_{Total} &= N(1-m)E_{Start} + NmE_{Start}(1+\alpha) \\ &= NE_{Start}(1+\alpha m) \end{aligned} \tag{5.7}$$

According to equation (5.7) there is (1+αm) times increment in the total energy. Normal nodes and advance nodes should have distinct cluster head election probabilities [20]. Weighted election probability for normal node can be given by

$$p_{nrm} = \frac{p_{opt}}{1+\alpha m} \tag{5.8}$$

And as the threshold value depends on election probability, the new threshold value for the normal node is:

$$T(n_{nrm}) = \begin{cases} \frac{p_{nrm}}{1-p_{nrm}\times(r mod \frac{1}{p_{nrm}})} & \text{if } n \in G' \\ 0 & \text{otherwise} \end{cases} \tag{5.9}$$

Where G‘ = set of normal nodes not selected as a cluster head in last 1/p_{nrm} round.

The value of election probability for advance nodes is evaluated as

$$p_{adv} = \frac{p_{opt}}{1+\alpha m} \times (1+\alpha) \tag{5.10}$$

And new threshold value for the advance nodes is:

$$T(n_{adv}) = \begin{cases} \frac{p_{adv}}{1-p_{adv}\times(r mod \frac{1}{p_{adv}})} & \text{if } n \in G'' \\ 0 & \text{otherwise} \end{cases} \tag{5.11}$$

Where G“ = set of advance nodes not selected as a cluster head in last 1/p_{adv} round.

Nodes select a random value according to equation (5.1). Three processes for cluster head selection LEACH, SEP and the proposed solution are compared in the table (5.4).

The proposed scheme selects high residual energy nodes A, B and G as the cluster heads for round r as it utilizes node heterogeneity for efficient operation.

The proposed scheme selects the high residual energy nodes as the cluster head for the next round and ignores the low residual energy nodes. Therefore high energy nodes act as the cluster head for first few rounds of the epoch. Thus high energy nodes are selected as cluster head in earlier rounds of an epoch. Selection of high energy nodes as cluster heads in the earlier rounds of an epoch does not suggest that the nodes left for the later rounds are inefficient because the nodes which are selected as cluster heads are the best suitable candidates for the job in that round . Energy saved in the earlier rounds is saved for the entire epoch and the rounds, if any,in which higher energy is dissipated will not effect the saved energy. Thereby proposed solution leads to well load balanced network

5.3 Performance Evaluation

The NS-2 simulator is used to implement and simulate the proposed scheme. To evaluate and compare the performance of the proposed scheme with LEACH and SEP following matrices are considered: *Network Lifetime* and *Data Units Received at Base Station.*

Table 5.4: Cluster Head Selection in LEACH, SEP and Proposed Solution for Heterogeneous network

Node	Residual Energy (Joule)	LEACH		SEP		Proposed Solution	
		Threshold	Random Number	Threshold	Random Number	Threshold	Random Number
A*	2.0	0.5	0.71	0.45	0.61	0.45	0.38
B*	1.5	0.5	0.5	0.45	0.42	0.45	0.29
C	1.0	0.5	0.63	0.6	0.54	0.6	0.67
D	0.5	0.5	0.42	0.6	0.71	0.6	0.93
E	1.2	0.5	0.64	0.6	0.78	0.6	0.61
F	0.4	0.5	0.4	0.6	0.35	0.6	0.77
G*	1.7	0.5	0.82	0.45	0.81	0.45	0.41

*(Advance nodes)

5.3.1 Initial Homogeneous Networks

Three topologies for varying number of nodes and size of the cluster are generated: 50×50 m^2 with 50 nodes, 100×100 m^2 with 100 nodes and 150×150 m^2 with 200 nodes. Nodes are distributed randomly over the application area and same network topologies are considered for all the simulated protocols. All nodes are homogeneous in nature with initial energy of 2 Joules. Packet size of header and data are 25 bytes and 500 bytes respectively. Radio energy model and network parameters used are described in chapter 3. Round time varied according to network dynamics as already discussed in chapter 4.

Figure 5.1 shows simulation time of first node death. We observe that SEP takes advantage, over LEACH, of node heterogeneity caused by network dynamics. Proposed scheme takes into account both residual node-energy and average network-energy for cluster head selection thereby increasing the probability of selecting a high residual energy node as a cluster head in earlier rounds of the network operations or epoch. The network is well load balanced resulting in extended first node death time over LEACH and SEP. Proposed scheme increases the first node death time by 31% and 45% over LEACH for 50×50 m^2 with 50 nodes and 100×100 m^2 with 100 nodes topology respectively. While, over SEP, an increment of 16% and 19% in the time to first node death is observed for 50×50 m^2 with 50 nodes and 100×100 m^2 with 100 nodes topologies respectively

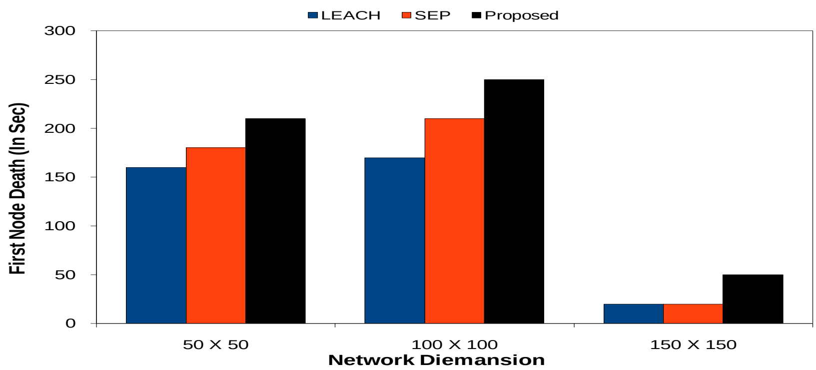


FIGURE 5.1: First Node Death for All Three Network Topologies

Figure 5.2 and 5.3 shows simulation time of 50% node death and 95% node death respectively. Proposed solution also prolongs the time of 50% node death and 50% node death. Time of 95% node death is extended by 20%, 15% and 10% by proposed solution over LEACH for 50×50 m^2 with 50 nodes, 100×100 m^2 with 100 nodes and 150×150 m^2 with 200 nodes topologies respectively. Time of 95% node death is extended by 11%, 11% and 7% by proposed solution over SEP for 50×50 m^2 with 50 nodes, 100×100 m^2 with 100 nodes and 150×150 m^2 with 200 nodes topologies respectively. LEACH and SEP do not take residual energy of a node into account while accessing the possibility for it to be selected as a cluster head. Proposed solution considers the residual energy of all the nodes before selecting a node as a cluster head and therefore improving the network lifetime over other two protocols.

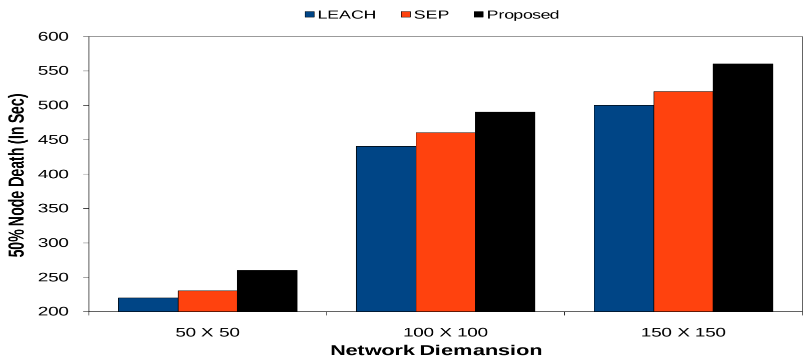


Figure 5.2: 50% Node Death for All Three Network Topologies

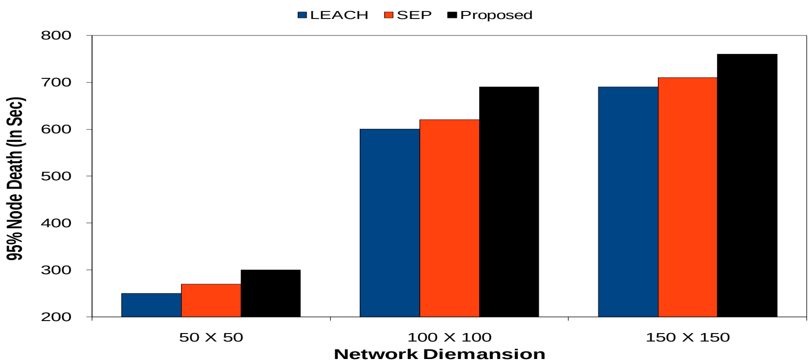


Figure 5.3: 95% Node Death for All Three Network Topologies

The number of data units successfully received at the base station for various networks are shown in figure 5.4. As the proposed solution has the advantage of increased network lifetime, it provides sensor nodes with more time to sense the surroundings and send the gathered information to base station. Proposed solution shows an improvement of 12%, 5% and 6% in data units received at base station for 50×50 m^2 with 50 nodes, 100×100 m^2 with 100 nodes and 150×150 m^2 with 200 nodes topologies respectively over LEACH. There is a significant improvement over SEP as well. In the proposed solution reliability of the networks is increased due because it is well load balanced.

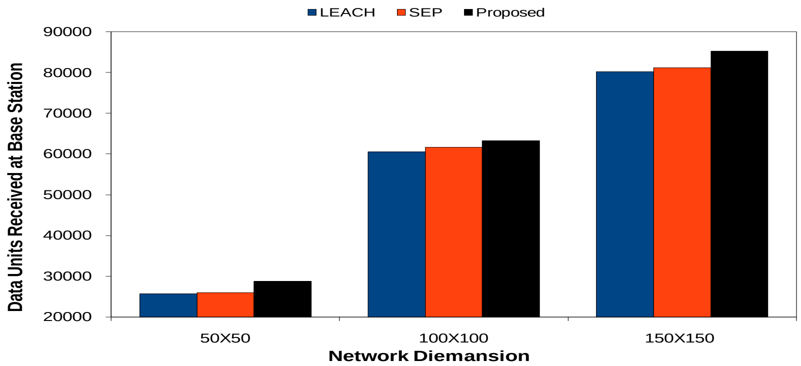


FIGURE 5.4: Data Units Received at Base Station for All Three Network Topologies

5.3.2 Initial Heterogeneous Networks

Two different network topologies, 50×50 m^2 with 50 nodes and 100×100 m^2 with 100 nodes, are generated for simulation of LEACH, SEP and proposed solution. Normal nodes are assigned an initial energy of 2 Joules whereas for initial energy of advance nodes (α) is varied from 1 to 5 (initial energy of advance nodes = α × initial energy of normal node). Percentage of advance nodes is set to be 10% and 20% of total nodes , i.e. Value of m is set to 10% and 20%. Heterogeneous nodes are also distributed randomly along with the normal nodes.

a. 10% nodes of Total Nodes that have More Energy in 50×50 m^2 with 50 Nodes Topology

Figure 5.5 and 5.6 show simulation time of first node death and 90% node death for different values of α. LEACH protocol does not utilize node heterogeneity hence time of first node death and 90% node death is almost same for all different values of α. SEP has different epochs for advance and normal nodes and is benefitted by node heterogeneity therefore time of first node death and 90% node death is better than LEACH and increases with the value of α. Proposed solution selects cluster head node according to residual energy and average energy of network so it takes advantage of node heterogeneity. As a result proposed solution has better network lifetime than that of LEACH and SEP.

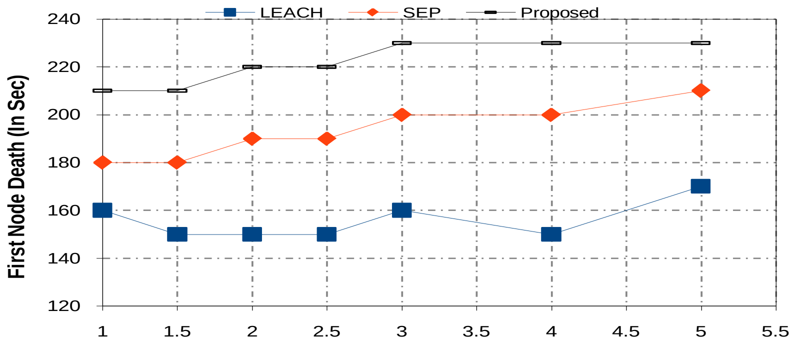


FIGURE 5.5: First Node Death Time for varying α

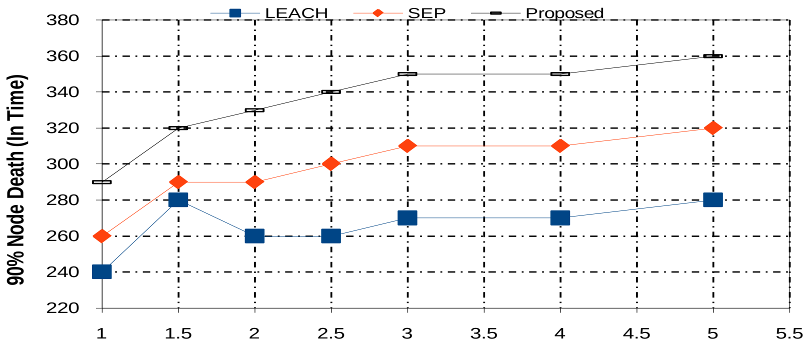


FIGURE 5.6: 90% Node Death Time for varying α

Figure 5.7 replicates the analysis of figures 5.5 and 5.6 for data units received at base station. Proposed scheme has remarkable improvement over LEACH and SEP for data units received at base station when 90% of the nodes are dead.

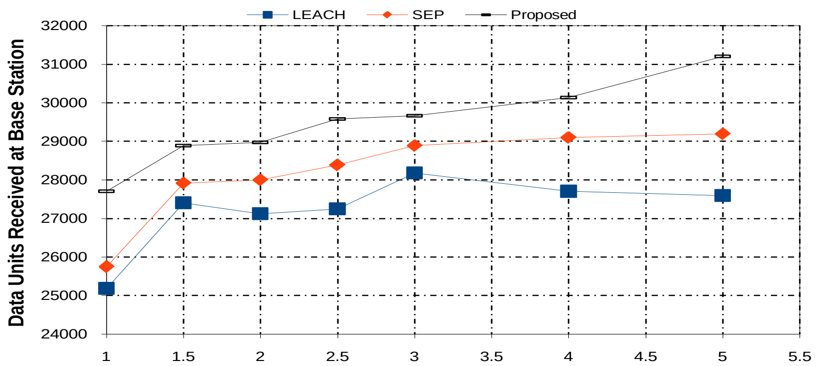


FIGURE 5.7: Data Units Received at Base Station for varying α

b. 20% nodes of Total Nodes that have More Energy in 50×50 m^2 with 50

Nodes Topology

Figure 5.8 and 5.9 show network lifetime for LEACH, SEP and proposed solution. 20% of the total nodes have extra energy that normal nodes, i.e. advance nodes in the network are 20% of the total nodes. It is analyzed from figure 5.8 and 5.9 that network lifetime increases with increase in node heterogeneity in case of LEACH also. Proposed solution is benefited by the increase in node heterogeneity, consequently increases lifetime of network consequential as compared to LEACH and SEP. We consider 80% node death time because there are 20% advance nodes.

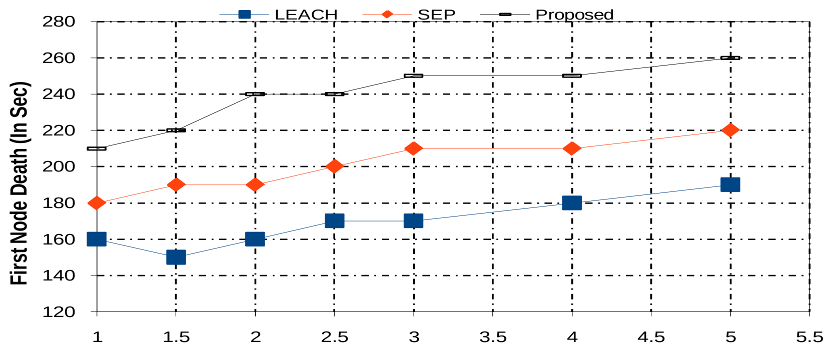


FIGURE 5.8: First Node Death Time for varying α

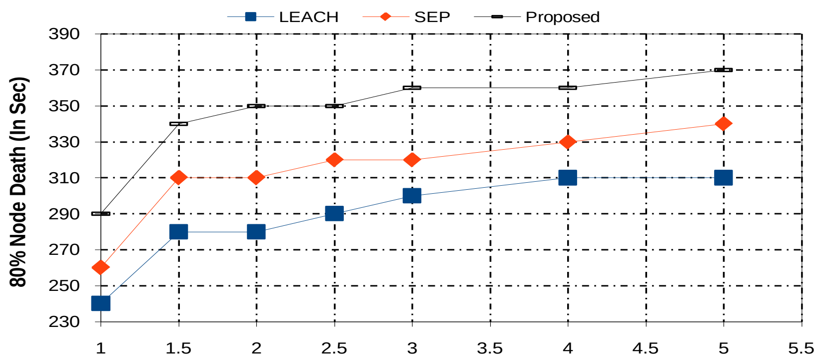


FIGURE 5.9: 80% Node Death Time for varying α

Figure 5.10 shows data units received at base station for various values of α at the time of 80% node death because it is expected that only advance nodes are alive in the later round (It is analyzed later in this section). As seen in figure 5.8 and 5.9, for proposed solution network lifetime is increased so nodes will send more data at base station. It is apparent from figure 5.10 that more data units are received at base station in case of proposed solution as opposed to LEACH and SEP protocols. Also, it shows that there is an increase in data units received at base station with the increase in initial energy of advance nodes.

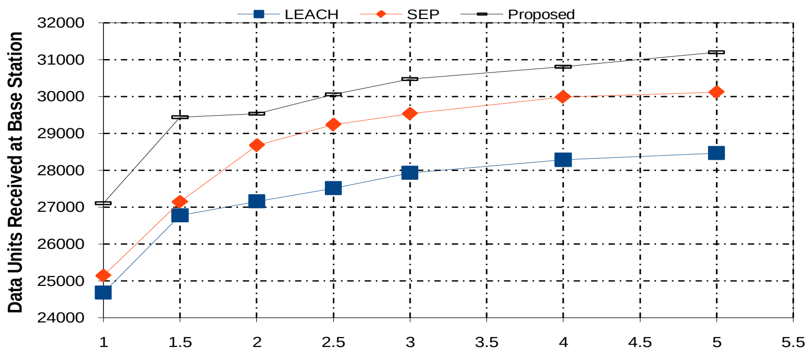


FIGURE 5.10: Data Units Received at Base Station for varying α

c. 10% nodes of Total Nodes that have More Energy in 100×100 m^2 with 100 Nodes Topology

Figure 5.11 and 5.12 show network lifetime for 100×100 m^2 with 100 nodes topology which has 10% of total nodes as advance nodes. LEACH protocol does some unknown advantage of node heterogeneity but it is outperformed by SEP and proposed solution. In case of proposed solution, network lifetime increases with the increase in energy of advance nodes. High energy advance nodes have more chances of being selected as cluster head and as the extra energy factor increases these chances also improve. Thus advance nodes are selected as cluster head more times as compared to normal nodes hence network is well load balanced in proposed solution. There is noticeable improvement in time for first node death and 90% node death with proposed solution over LEACH and SEP. Figure 5.13 shows data units received at base station for three protocols. Proposed solution has more data units received at base station as compared to LEACH and SEP.

d. 20% nodes of Total Nodes that have More Energy in 100×100 m^2 with 100 Nodes Topology

Simulation time for first node death and 80% node death are depicted in Figure 5.14 and 5.15 for different values of α. Performance of LEACH improves with the increase in the value of α, i.e., percentage of advance nodes. But proposed scheme makes better use of the node heterogeneity in order to lengthen the network lifetime. As can be seen

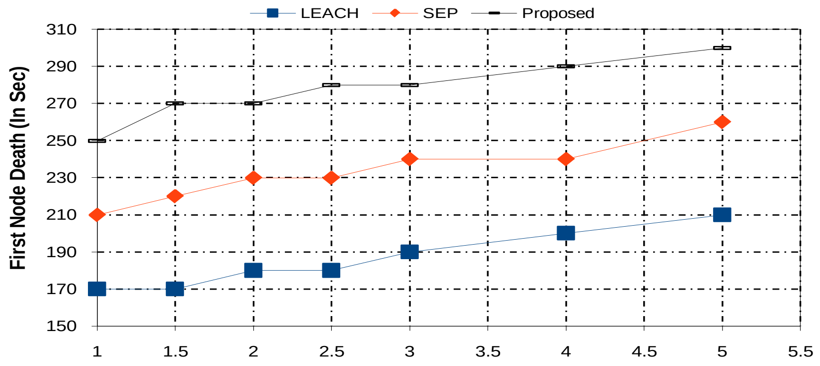


FIGURE 5.11: First Node Death Time for varying α

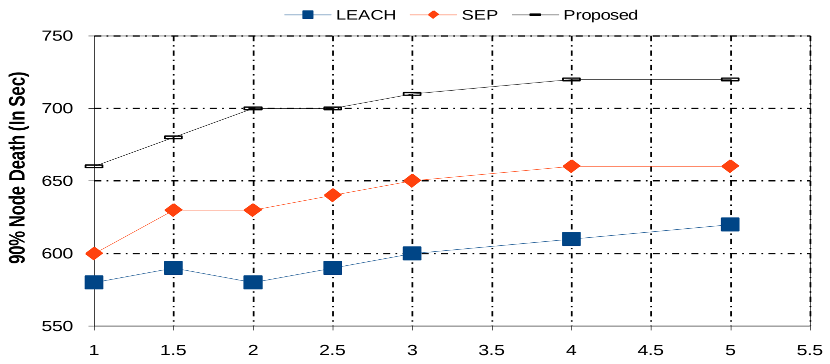


FIGURE 5.12: 90% Node Death Time for varying α

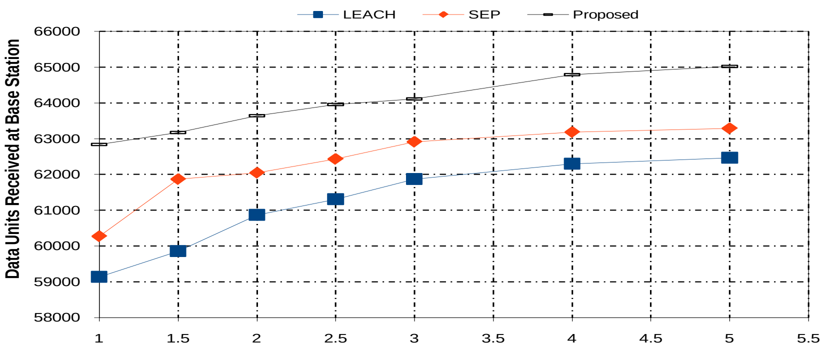


FIGURE 5.13: Data Units Received at Base Station for varying α

from figure 5.16, data units received at base station are also significantly improved by proposed solution.

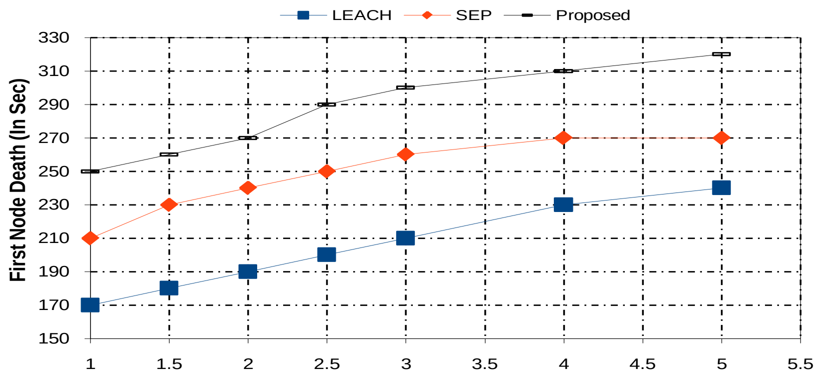


FIGURE 5.14: First Node Death Time for varying α

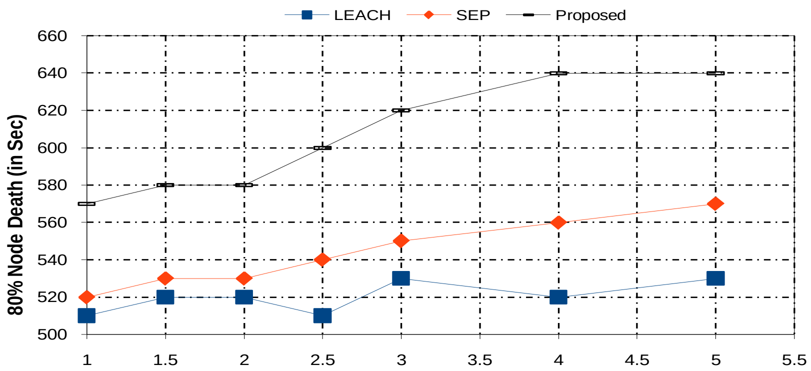


FIGURE 5.15: 80% Node Death Time for varying α

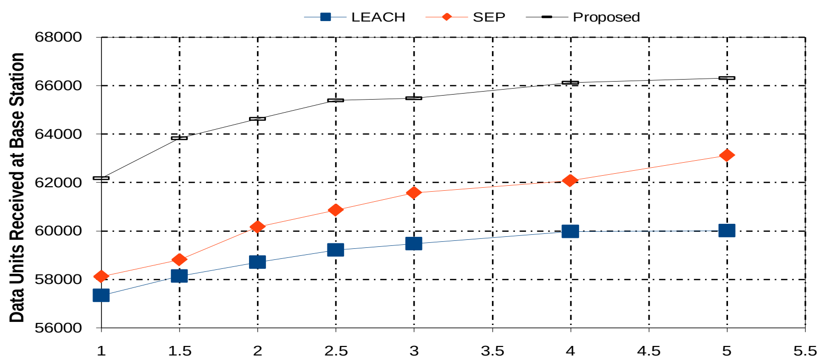


FIGURE 5.16: Data Units Received at Base Station for varying α

5.4 Conclusion

In this chapter, an improved and efficient cluster head selection solution, which is capable to adapt the heterogeneous wireless sensor network environment, is proposed. Selection of a node as a cluster head is based on the criteria decided by taking into consideration both residual node energy and average energy of network at beginning of any particular round. High residual energy nodes are preferred to lower residual energy nodes which leads to better load balancing of the network. As evident from simulation results, proposed clustering algorithm provides longer network lifetime and data gathering in contrast to the other prevalent scheme LEACH and SEP.

Chapter 6

Balance-Sized Cluster Solution to Extend Lifetime of Wireless Sensor Networks

Traditional clustering algorithms have uneven clusters that make network load unbalanced. In this Chapter, a clustering approach is proposed that yields balanced clusters by considering thresholds for cluster formation. Simulation results prove that proposed scheme provides increased network lifetime and lower node death rate compared to traditional clustering algorithm.

6.1 Introduction

Lack of human access because of harsh/remote areas makes it virtually impossible to recharge or replace a battery of sensor nodes. Therefore, efficient energy consumption of nodes is prime design issue for wireless sensor networks. One such approach is clustering which is energy efficient approach for wireless sensor networks to extend network lifetime. Reduced communication distance, TDMA schedule and data aggregation/fusion in clustering algorithm save energy of nodes. Clustering algorithms make network load balanced and scalable. In [5, 24], author points importance of equal size clusters. But most of existing clustering algorithms have limitation of uneven cluster sizes. Clusters of small and large size exist for same time phase. It makes network load unbalanced and hence, thereby decreasing the network lifetime eventually.

In this work, we propose a balanced cluster size clustering approach to protract the network lifetime. Proposed solution has threshold for number of nodes in clusters ($Th_{cluster}$) for initial cluster formation and distance threshold ($Th_{distance}$) for un-clustered nodes to join cluster. Un-clustered nodes, after initial cluster formation, join cluster head according to $Th_{distance}$ and $Th_{cluster}$. Proposed solution has better balanced cluster formation depending upon thresholds and hence achieves longer network lifetime and lower node death rate compared to traditional clustering algorithm. Remaining chapter is sectioned as follows. Section 6.2 describes proposed method and section 6.3 compares cluster formation and performance of proposed method with existing clustering algorithm. Finally, section 6.4 concludes work of chapter.

6.2 Proposed Solution

6.2.1 Problem Statement

In most of the existing clustering algorithms, phase of data transmission is segregated into frames and frames are further broken into time slots. Each node in cluster is then assigned a time slot in each frame. After each frame duration, cluster head aggregates the data and transmits it to base station. So, length of frame depends upon number of member nodes in a cluster. If a cluster has n nodes, T_{trans} is time needed for transmission of data packet and T_{DA} is the total time cluster head takes to aggregate data and transmit it to base station. Then, frame length (time) can be calculated as:

$$Frame\ Time = (n \times T_{trans}) + T_{DA} \tag{6.1}$$

If T_{round} is the time of data transmission phase, then number of time slots available for a node in a round are:

$$\begin{aligned} Time\ Slots\ per\ node &= \frac{T_{round}}{Frame\ Time} \\ &= \frac{T_{round}}{(n \times T_{trans}) + T_{DA}} \end{aligned} \tag{6.2}$$

According to (6.2), nodes of small sized clusters have more time slots in data transmission phase, so consume high energy as compared to nodes of large sized clusters. Further, head of small sized cluster consumes more energy because it sends data more times to base station compared to head of large cluster. Consider two clusters , A and B, have members nodes 10 and 20 respectively. T_{trans} is considered 0.5 sec, T_{round} is 20 sec, and energy consumption in radio communication is as in [51]. Cluster head of A consumed energy 0.35J while 0.29J is consumed by cluster head of B. Average energy consumption of member node of A and B is 0.021J and 0.016J respectively. Small clusters over-sensed a region while area under large clusters are under-sensed over time. So a better uniform cluster formation approach is required to have well load balanced network.

6.2.2 Proposed Clustering Algorithm

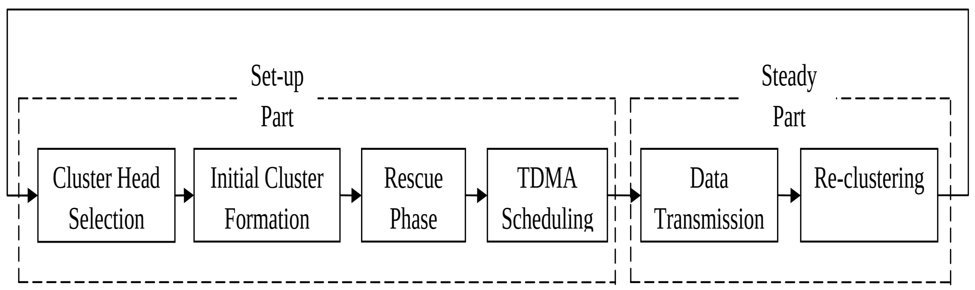


FIGURE 6.1: Operation of Proposed Method

In this chapter, a distributed clustering approach is presented to have balanced clusters. Operation of proposed solution is shown in figure 6.1. Selected cluster heads broadcast status message to network. Advertisement of selected cluster heads is received by other heads as well, so all the heads have the information of number of cluster heads(X) present in the network. In the proposed algorithm, cluster formation is carried out in two phases: *Initial cluster formation* and *Rescue Phase*. In the initial cluster formation, $Th_{cluster}$ defines number of member nodes in each cluster. Due to distributed and random nature, cluster heads selected in each round are different, so $Th_{cluster}$ is calculated for each round. If there are N active nodes in network then $Th_{cluster}$ is:

$$Th_{cluster} = \frac{N}{X} \tag{6.3}$$

Nodes join nearest cluster head according to RSS of advertisements. If selected head has member nodes less than Th$_{cluster}$, then node joins that head otherwise it waits for the rescue phase. At the end of this phase, a cluster has number of nodes equal to or less than the Th$_{cluster}$. In rescue phase, un-clustered nodes join the best possible cluster. If a head is placed at a distance less than Th$_{distance}$ from an un-clustered node and has member nodes less than Th$_{cluster}$, then the un-clustered node will join that cluster. Node will join nearest cluster head if the joining condition of rescue phase is not satisfied by any cluster head. Th$_{distance}$ is calculated considering the tradeoff between cluster size and total cluster distance for better cluster quality. In the next section, an explanation for the selection of the value of Th$_{distance}$ has been given. Algorithm 1 presents the rescue phase.

Algorithm 1 Rescue Phase

Input: Number of CHs(X), Distance of node i to CHs(DIST matrix), Number of nodes in each cluster(CLUSTER matrix)
Output: Cluster Head Selected

```
Arrange DIST matrix of node i in increasing order, CLUSTER matric with corre-
sponds of DIST
JOIN=1
for j:=2 to X do
    if DIST[j] ≤ Th_distance then
        if CLUSTER[j] < Th_cluster then
            JOIN=j
        end if
    end if
end for
node i selects cluster head JOIN.
```

Distance between the two communicating nodes is the prime attribute for energy consumption of the nodes. So, in initial cluster formation phase nodes join nearest cluster head. The remaining un-clustered nodes join the cluster head according to Th$_{cluster}$ and Th$_{distance}$. Again the distance between the node and cluster head is considered in the join condition.

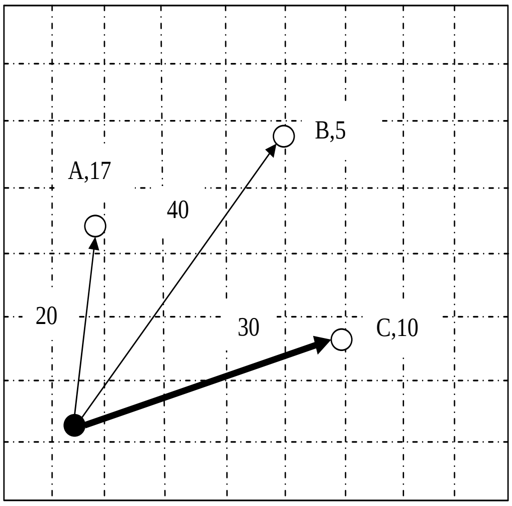
Figure 6.2(a-c) shows different possibilities of join for a node. Consider a network of 50×50 m^2 with 50 nodes and there are three cluster heads, A, B and C. So Th$_{cluster}$ is 17 and we set Th$_{distance}$ as 40m. Dark circle shows un-clustered node and circle shows cluster head node. Distance of cluster heads from the node is marked. Name of the cluster head along with the number of nodes in the cluster is shown, i.e. A,17 means cluster head A has 17 member nodes. Dark line shows decided cluster head for join after execution of algorithm 1. In case-1, cluster heads A, B and C have 17, 5 and 10 nodes respectively after initial cluster formation phase. Cluster head A is nearest to node but already has member nodes equal to $\mathrm{Th}_{cluster}$ so node will look for other best possible cluster head. Cluster head C is closer than B and also has member nodes less than $\mathrm{Th}_{cluster}$. So node will join cluster head C in that case. In case-2, node will not join cluster head A and C as both cluster heads have member nodes equal to $\mathrm{Th}_{cluster}$. Cluster head B is placed less than $\mathrm{Th}_{distance}$ and have member nodes less than $\mathrm{Th}_{distance}$. Therefore, node join cluster head B in that case. Case-3 is different from above cases. Cluster head A is nearest but has member nodes equal to $\mathrm{Th}_{cluster}$, cluster heads B and C have member node less than $\mathrm{Th}_{cluster}$ but both have distance greater than $\mathrm{Th}_{distance}$. None of cluster heads fulfill the join condition of rescue phase, so node will join the nearest cluster head. In that case, node joins cluster head A. Effect of $\mathrm{Th}_{distance}$ on cluster quality is shown in section 6.3.1.

Cluster heads are selected based on probabilistic approach. Each node selects a random number between 0 and 1. T(n), threshold is calculated that depends upon current round(r) and optimal cluster head probability(P_{opt}). If selected random number falls below threshold for a particular node then that node is selected as cluster head.

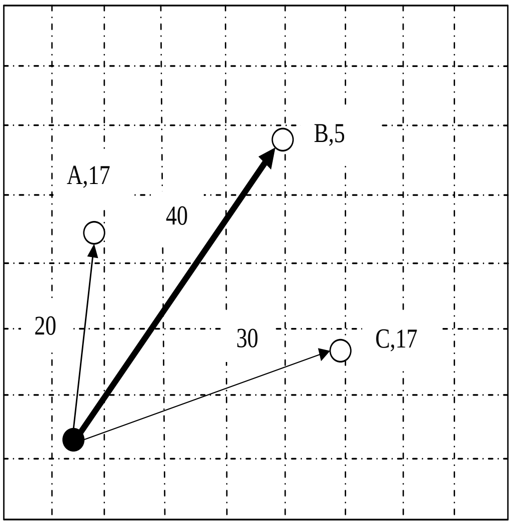
$$T(n) = \begin{cases} \frac{P_{opt}}{1-P_{opt}\times(r mod \frac{1}{P_{opt}})} & \text{if } n \in G \\ 0 & \text{otherwise} \end{cases} \tag{6.4}$$

where G = set of nodes that are not selected as cluster head in last $1/\mathrm{P}_{opt}$ rounds. If selected number is less than threshold then node is selected as cluster head. Proposed solution performs operation in rounds and nodes are selected as cluster head in a round according to probabilistic approach which ensures that each node is selected as cluster head once in an epoch of $1/\mathrm{P}_{opt}$ rounds.

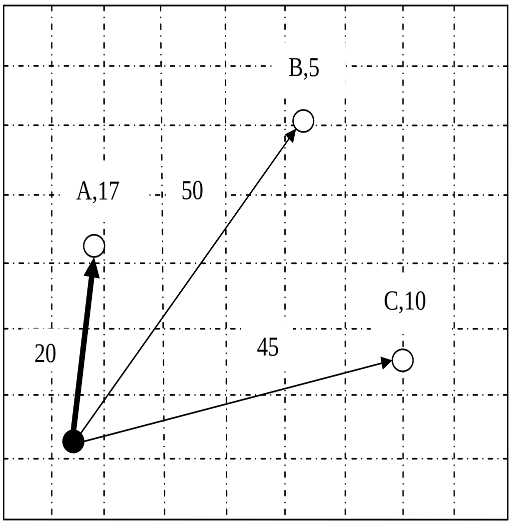
After cluster set-up, heads constitute TDMA schedule and broadcast to member nodes. Nodes wake up to send the data to cluster head as time slot arrives otherwise remain in sleep state for rest of the frame time. After completion of current round, re-clustering is done for next round.



(a) case-1



(b) case-2



(c) case-3

FIGURE 6.2: Rescue Phase

6.3 Performance Evaluation

Network of 50×50 m^2 with 50 nodes, base station located outside (25,100) is simulated in MATLAB. All nodes are homogeneous and initial energy of nodes is 2.0J. Radio energy model described in [51] is used.

6.3.1 Effect of $Th_{distance}$ on Cluster Quality

Cluster quality shows how efficient are clusters in terms of size and total cluster distance. Number of sensor nodes in a cluster is the size of that cluster. Sum of the distances of member nodes to their respective cluster head is total cluster distance and is calculated as:

$$Total\ Cluster\ Distance = \sum_{i=1}^{X}\sum_{j=1}^{n} Dist(j,i) \tag{6.5}$$

where X is number of CHs, n is number of nodes in a cluster and distance of member nodes to respective CH is given by function Dist ( ).

As can be analyzed from algorithm 6.1, un-clustered nodes may join cluster head that is placed far from the nearest cluster head if join condition is satisfied or may join the nearest cluster head if join condition is not satisfied by any of the cluster heads. So, the total cluster distance may increase for proposed solution that depends upon $Th_{distance}$. Size of clusters after rescue phase also depends on the same. We varied the $Th_{distance}$ and analyze the effect on cluster quality. Figure 6.3 shows cluster formation in LEACH. Figures 6.4, 6.5, 6.6 and 6.7 show cluster formation for proposed solution for $Th_{distance}$ of 30m, 40m, 50m, and 60m respectively. As seen from these figures, with the increase in value of $Th_{distance}$ clusters are more balanced. LEACH has very uneven cluster formation and there are almost equal-sized clusters in each run for $Th_{distance}$=60m. Figure 6.8 compares total cluster distance for proposed solution of different values with LEACH. Balanced-sized clusters come with the cost of increased total cluster distance. With the increase in $Th_{distance}$, most of the un-clustered nodes satisfy the join condition of proposed solution so have less probability of joining the nearest cluster head. So there is increase in total cluster distance. Cluster quality increases for cluster size while it decreases for total cluster distance for proposed solution with the increase in $Th_{distance}$.

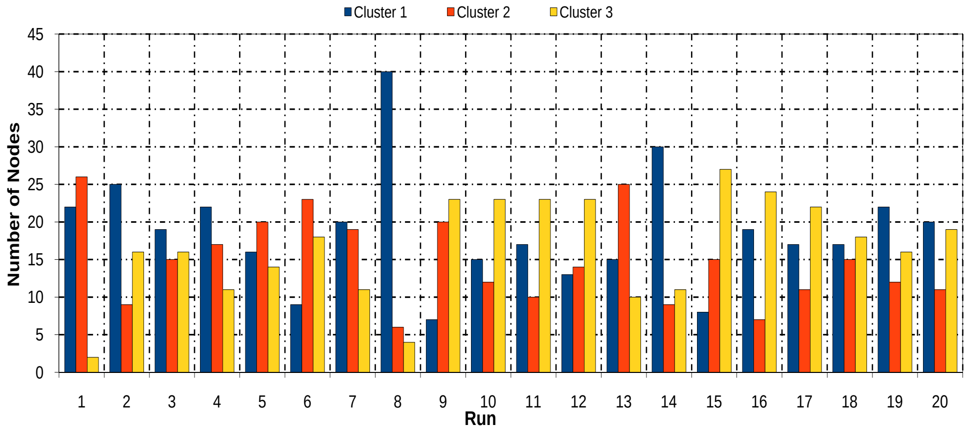


FIGURE 6.3: Clusters formation in LEACH

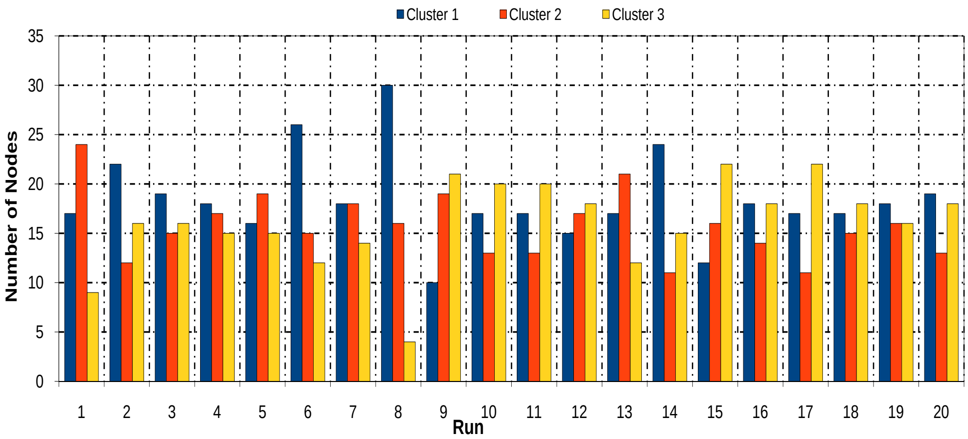


FIGURE 6.4: Clusters formation for $Th_{distance}$=30m

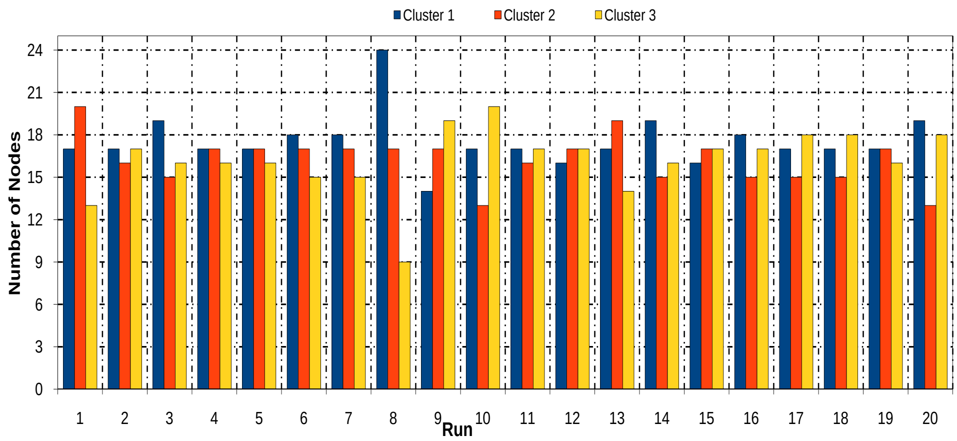


FIGURE 6.5: Clusters formation for $Th_{distance}$=40m

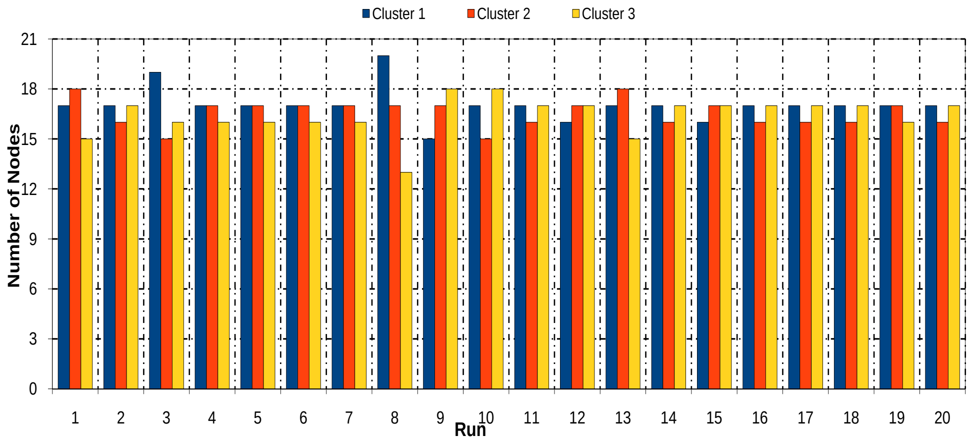


FIGURE 6.6: Clusters formation for $Th_{distance}$=50m

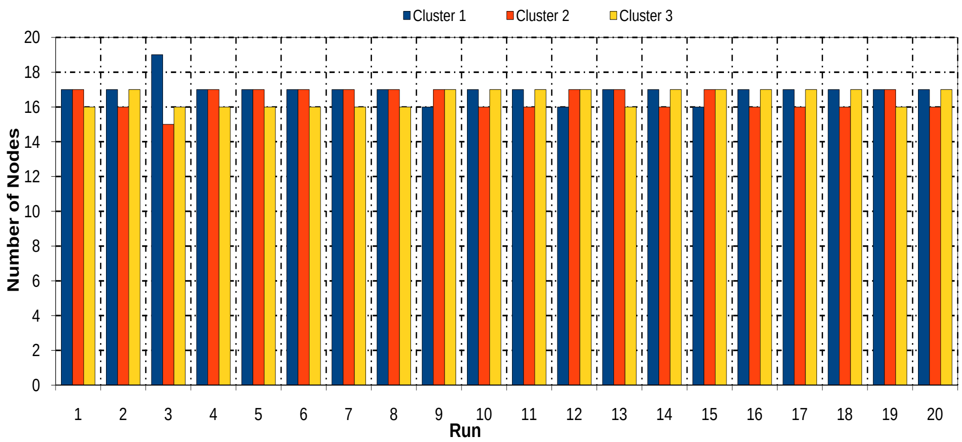


FIGURE 6.7: Clusters formation for $Th_{distance}$=60m

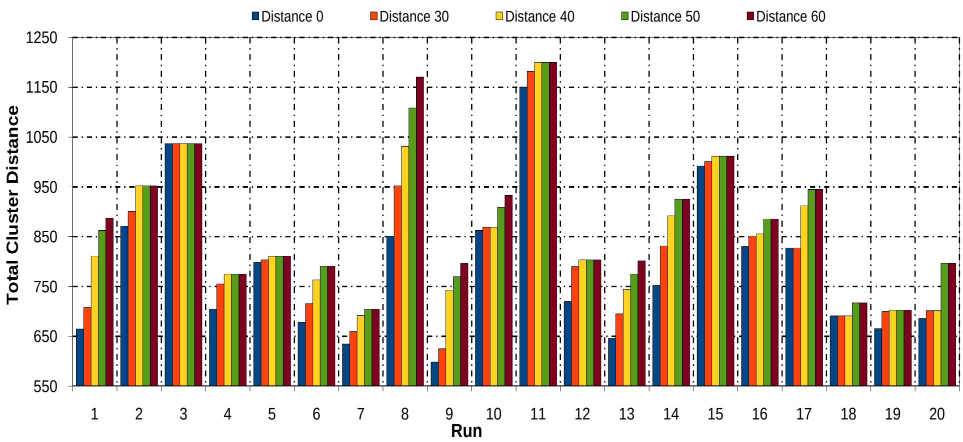


FIGURE 6.8: Total Cluster Distance

Table 6.1 combines analysis of above figures and shows effect of $\mathrm{Th}_{distance}$ on cluster quality. With increase in $\mathrm{Th}_{distance}$, clusters are more balanced but with the cost of increased total distance. So, considering trade-off between cluster size and total distance, we set $\mathrm{Th}_{distance}$ as 40m to have better cluster quality for our work.

Table 6.1: Comparison of Cluster Quality (Cluster Sizes and Total Cluster Distance)

		LEACH	**30m**	**40m**	**50m**	**60m**
Run 1	Cluster1	22	17	17	17	17
	Cluster2	26	24	20	18	17
	Cluster3	2	9	13	15	16
	Distance(m)	665	708	801	870	887
Run 2	Cluster1	25	22	17	17	17
	Cluster2	9	12	16	16	16
	Cluster3	16	16	17	17	17
	Distance(m)	871	901	952	952	952
Run 3	Cluster1	40	30	24	20	17
	Cluster 2	6	16	17	17	17
	Cluster3	4	4	9	13	16
	Distance(m)	851	932	1031	1108	1170
Run 4	Cluster1	19	19	19	19	19
	Cluster2	15	15	15	15	15
	Cluster3	16	16	16	16	16
	Distance(m)	1036	1036	1036	1036	1036

6.3.2 Experimental Results

For analysis and comparison of proposed algorithm to traditional clustering algorithm, *Node death rate*, *network lifetime* and *energy consumption of network over rounds* is considered.

Figure 6.9 shows number of nodes alive against number of rounds. As seen, there is larger stable region for proposed method compared to LEACH. Number of nodes alive over rounds is always higher for proposed method compared to LEACH, i.e. node death rate of proposed solution is lower than LEACH, because energy consumption of nodes throughout the network is more uniform in proposed method, i.e. the network is more load balanced in proposed method compared to LEACH, (even though there is increase in overhead for cluster formation in proposed method).

As analyzed from figure 6.9, first node death in proposed method appears later than LEACH. There is also significant improvement for 50% node death and overall network

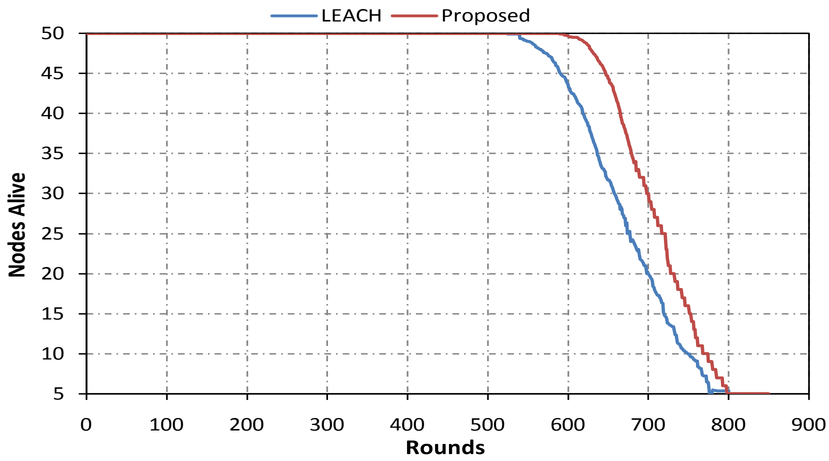


FIGURE 6.9: Node Death Rate

lifetime as shown in figure 6.10. There is increase of 13%, 8% and 6% for first node death, 50% node death and 95% node death for proposed solution over LEACH respectively.

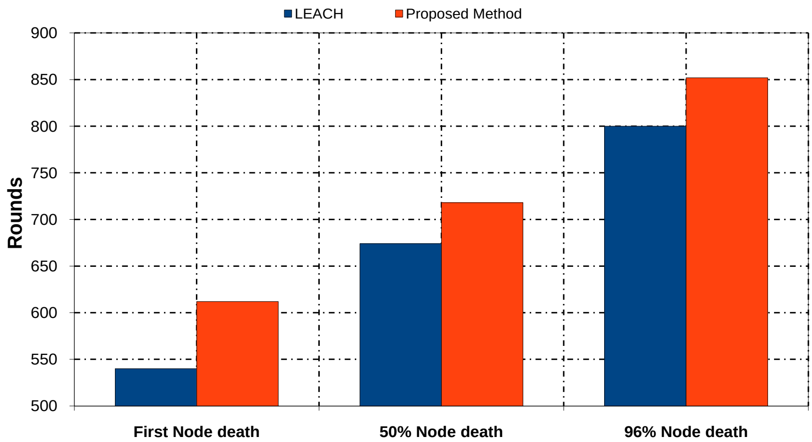


FIGURE 6.10: Network Lifetime

Lower node death rate and prolonged lifetime is achieved in proposed method over LEACH. There are balance-sized clusters in proposed solution so energy consumption is more uniform compared to LEACH. Figure 6.11 shows comparison of energy consumption of network over rounds for proposed solution and LEACH. Energy consumption for proposed solution is always less than LEACH.

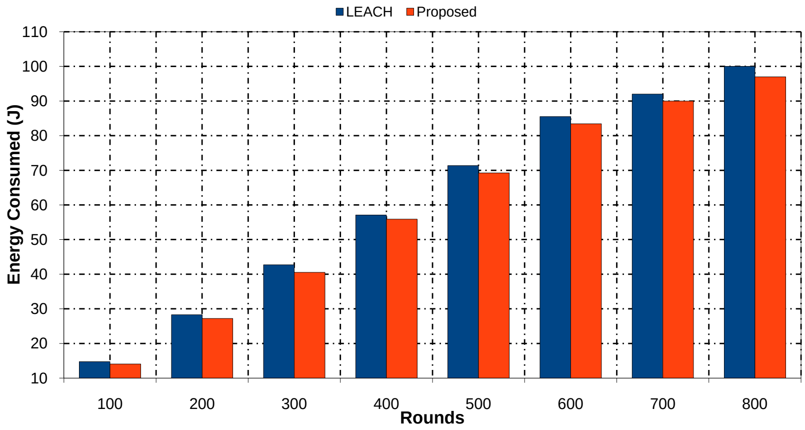


FIGURE 6.11: Energy Consumption of Network

6.4 Conclusion

Uneven size clusters exist simultaneously in clustering algorithms. Small size cluster has more frames in the data transmission phase than large cluster so energy consumption of small clusters is more than large size clusters that leads to load unbalanced network. Work of this chapter presented a clustering approach that have balanced-size clusters. Cluster formation process is divided into two phases: initial cluster formation and rescue phase. Initial cluster formation phase has threshold for maximum number of member nodes in a cluster. Few nodes are left which do not have any cluster head. Un-clustered nodes select cluster head according to join condition of rescue algorithm. Proposed method has balanced clusters and better cluster quality. Simulation results show that proposed solution outperform existing clustering algorithm and has extended network lifetime and low node death rate.

Chapter 7

Analyzing The Reliability State of Wireless Sensor Network using Genetic Algorithm

The reliability of wireless sensor network can be defined in terms of area covered by sensor nodes and redundancy in the sensed data. Redundancy in the data is originated by overlapping in the sensed area of nodes. It is necessary to know that the deployed wireless sensor network is reliable or not. As nodes start dying, area is not completely sensed and the redundancy of the data also decreases. Hence the gathered data from the network is also unreliable. So it is imperative to find when the deployed reliable network switches to unreliable state so that a proper action can be taken. In this chapter, genetic algorithm based method is proposed to find whether the deployed wireless sensor network is reliable or not.

7.1 Introduction

Wireless sensor networks are application specific and comprise of large number of sensor nodes. Wireless sensor networks can be categorized into two classes according to their applications: *Data Gathering* and *Event Driven*. In a data gathering network, nodes continuously sense the area and send the data consistently to base station. In an event-driven wireless sensor network, nodes send the data only when an event is triggered. In both types of networks, the information of area is send to base station.

Quality of the data gathered by the wireless sensor network is of principle concern. High quality of data can be gathered only from a reliable wireless sensor network. Reliability of wireless sensor network can be defined in terms of the area sensed by all nodes and the redundancy in the gathered data. Redundancy of data can be caused by either by sending information of a location by multiple sensors (spatial redundancy) or a sensor sends same information several times (temporal redundancy). Temporal redundancy causes additive energy consumption of network because energy is consumed in communication of data. Spatial redundancy is caused by the overlapped sensed area of nodes, consequently information of overlapped area is send by all these nodes so energy is consumed to send same piece of information. But, spatial redundancy is required to provide fault tolerance, to improve the reliability of the collected data and to provide information security [7]. As there is a dense deployment of sensor nodes in the field so there is always an overlapping in the sensed area of nodes. Hence, a network can be considered reliable if it senses the almost whole area of interest and produces the desired minimum redundancy in the sensed data for accurate information about the phenomenon.

Even after the implementation of energy efficient algorithms to have load balanced network, nodes go out of energy randomly. Sensor nodes consume inconsistent energy in the field and hence die randomly. As the number of nodes start decreasing in the field, the reliability of the network in terms of the covered (sensed) area and redundancy in the sensed data also starts decreasing. At that point, there will be blind points, regions that are not monitored by any sensor, in the network. The data gathered at that time is of low quality. So, it is necessary to find the state of wireless sensor network, whether it is reliable or unreliable.

Finding the minimum number of nodes (random) that covers the almost whole area with minimum desired overlapped area is NP-HARD issue [112]. Genetic algorithms [76] are used for many NP-HARD problems like optimizations and Traveling Salesman Problem (TSP). In this chapter, the problem of finding the minimum number of nodes (random) that covers the almost whole area with the desired minimum redundancy according to the application of the network is optimized by genetic algorithm (GA). Work of this chapter, analyzes the reliability of network considering only spatial redundancy. Rest of the chapter is organized as follows: section 7.2 defines the problem description. Section 7.3 gives the description of genetic algorithm for defined problem. Section 7.4 describes the experimental results and usefulness of the procedure and section 7.5 concludes the work.

7.2 Problem Description

Wireless sensor networks are application specific networks so the performance of the network should stick to the requirements of application. Wireless Sensor networks should be reliable enough to provide high quality information about the phenomenon. The two main requirements of wireless sensor networks to be considered as reliable are:

- Almost each point is covered, (i.e. all area is sensed)
- Minimum desired overlapped area, (i.e. amount of redundancy in the sensed data)

Sensor nodes are positioned in a spatial region and hence can be considered as points in two dimensional planes. If the positions of the sensors are pre-engineered, the minimum number of active nodes that meet the above stated requirements of a reliable network can be calculated as:

$$\frac{((Area\,of\,field) + (Minimum\,overlapped\,area))}{Area\,covered\,by\,a\,single\,node} \tag{7.1}$$

But in most of the application, sensor nodes are deployed randomly not by a pre-engineered procedure to engineer the location of nodes. Communication between the nodes is the prime source of energy dissipation of nodes that depends on the distance between the two. Energy consumption of nodes is not uniform and that makes the sensors nodes die randomly in the field. As the network progresses, the number of nodes alive in field starts decreasing. The reliability of the network also decreases as the decreased number of nodes will not provide enough redundant data. Minimum desired redundancy in the data depends on the application of network. Over the time, the nodes will not be able to sense the complete area along with redundancy below the desired level. The gathered data is not of good quality and does not provide accurate information about the phenomenon. Now the network can be considered in unreliable state.

Hence it is necessary to find, when the network will transit from a reliable state to unreliable state. So, find the minimum number of nodes (random) that makes the network reliable.

7.3 Genetic Algorithm

The problem stated in the above section becomes more complex due to the dynamic nature of sensor networks. Whereas, genetic algorithms are very flexible in solving such dynamic problems. In this work, genetic algorithm is applied in a way to find the constraint on the minimum number of nodes alive in the field while fulfilling the requirements for a deployed network.

GA maintains a population of chromosomes and each chromosome represents a solution. Each chromosome is evaluated to determine the fitness according to the fitness function that defines the problem. Genetic transformations, crossover and mutation, are applied to selected chromosomes. A new population is generated with the combination of chromosomes with better fitness in the current population and the new chromosomes generated by genetic transformation. After several generations, the algorithm converges to the best solution. Figure 7.1 shows the operation of genetic algorithm.

7.3.1 Genetic Algorithm for the Stated Problem

a. Population: The initial population comprises of randomly generated set of chromosomes. Binary representation is used for chromosome definition and each bit corresponds to one sensor node. So, the length of each chromosome is equal to the initial number of nodes in the field. Presence of a node in the field constitutes a “1” in the chromosome otherwise a “0”.

b. Fitness Function: Survival of a chromosome depends on its fitness. Fitness of each individual of the population is calculated. The problem detaled in the chapter has three parameters for calculating the fitness of chromosomes.

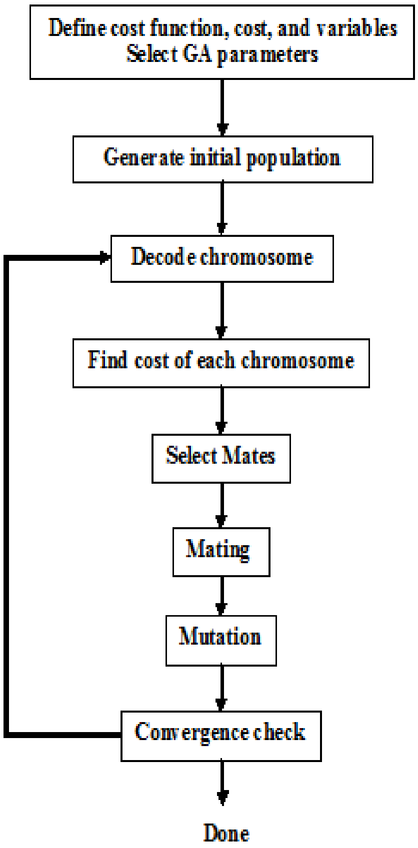


FIGURE 7.1: Genetic Algorithm

1. Number of active nodes should be minimum.
2. Nodes present in the field should cover almost the whole area.
3. Sensed area of nodes should overlap at least to the desired percentage.

If the number of active nodes is less than the minimum required nodes, that means, the network is not reliable any more. The network either not covers the whole area or the data produced does not have the minimum redundancy to make the data reliable.

So the fitness function for the problem consists of three objectives and is represented as:

$$Fitness = f(n, covered\ area, overlapped\ area) \tag{7.2}$$

After scaling the fitness function,

$$Fitness = \begin{cases} \frac{N-Active\ nodes}{N} + \frac{Total\ covered\ area}{N} \\ -\frac{\pm(overlapped\ area - Minimum\ overlapped\ area)}{Minimum\ overlapped\ area} \end{cases} \tag{7.3}$$

Fitness scaling is applied to fitness function for better selection procedure and to avoid premature convergence and slow finishing.

c. Selection: It is the process of selecting two parents from the population for crossing. The purpose of selection process in a genetic algorithm is to give more reproductive chances to those population members that are better fit. The selection procedure may be implemented in a number of ways like Roulette Wheel selection, Tournament selection, Boltzmann selection, Rank selection, Random selection, etc. In this work, Roulette Wheel selection procedure is applied to select chromosomes for generating new population. The procedure creates a biased roulette wheel in which the slots are sized in proportion to the fitness of chromosomes. Selection procedure is random but the chance of being selected is proportional to the fitness of chromosomes. Chromosomes having higher fitness values are more liable to be selected for the population of the next generation.

Due to the random nature of selection procedure, the best member of the population may not succeed to live on to the next generation. The elitist strategy fixes this loss by advancing the best member of current generation to the next generation. It improves the performance of genetic algorithm.

d. Crossover Operator: One-point crossover method is used in this work. The probability with which the crossover operation takes place between the two chromosomes is specified by the Crossover rate. The crossover point separates the portions exchanged by the two chromosome. The example of one point crossover is as follows:

Chromosome 1	0	1	1	1	0	**0**	**1**	**1**	**1**	**0**	**1**	**0**
Chromosome 2	1	0	1	0	1	**1**	**0**	**0**	**1**	**0**	**1**	**0**

This crossover engenders two offspring given as below:

Offspring 1	0	1	1	1	0	**1**	**0**	**0**	**1**	**0**	**1**	**0**
Offspring 2	1	0	1	0	1	**0**	**1**	**1**	**1**	**0**	**1**	**0**

e. Mutation Operator: Each bit of a chromosome is operated by the mutation operator with a probability of mutation rate resulting in selected bit reversal.

Before	0	1	1	**1**	0	0	1	1	1	0	1	0
After	0	1	1	**0**	0	0	1	1	1	0	1	0

7.4 Results

The purpose of this work is to optimize the problem explained in section 7.2 using genetic algorithm. Sample sensor networks are generated with different number of nodes deployed over different dimensions. The node density of all sample networks is almost the same. All nodes are stationary once deployed in the field. Parameters used are listed in Table 7.1.

Table 7.1: Parameters and Values

Parameters	**Values**
Number of Nodes(N)	25,100,150
Sensing Range	10m
Network Dimension	50×50 m^2, 100×100 m^2, 125×125 m^2
Size of Population	N
Length of Chromosome	N
Selection	Roulette Wheel
Crossover Rate	0.65-0.75
Type of Crossover	one-point
Mutation Rate	0.01-0.02

For our work, the minimum acceptable overlapping in the sensed area is varied from 20% to 30% of the whole area.

7.4.1 For Minimum 20% Overlapped Area

a. For 50×50 m^2 with 25 Nodes: Figure 7.2 shows the best fitness value over the generations. There is an increase in the best fitness value of the current population over generations. After 170 generations the best fitness value is constant, .i.e. genetic algorithm applied optimizes the problem. As the fitness function is increasing the value of the best fitness value so the number of active nodes is decreasing to an optimal value as shown in figure 7.3. The results show that the solution for the stated problem converges at 12 nodes (random), i.e., 48% of the initial nodes in the field.

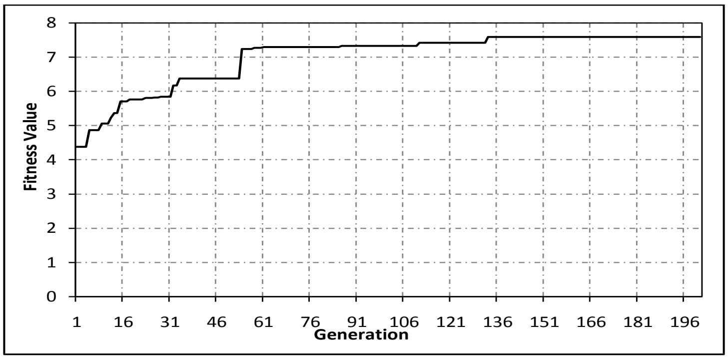


Figure 7.2: Best Fitness Value Vs Generation

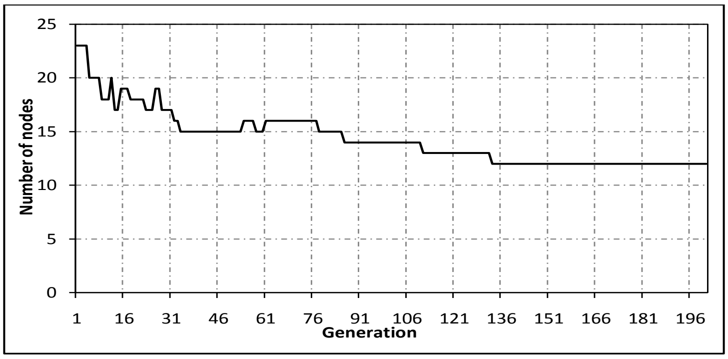


Figure 7.3: Active Nodes Vs Generation

b. For 100×100 m^2 Area with 100 Nodes: Figure 7.4 shows that the fitness function in equation (7.2) increases the best fitness value of the population over generations. The applied genetic algorithm approach has the constant best fitness value after 250 generations. Figure 7.5 shows that the number of active nodes is decreasing over the generation and the solution for the stated problem is converging at 42 nodes (random), i.e., 42% of the initial nodes in the field.

c. For 125×125 m^2 Area with 150 Nodes: Figure 7.6 shows the increase in the best fitness value over generations. After 300 generations the increase is slight and after 470 generations it converges. Figure 7.7 shows that the solution for the stated problem converges at 64 nodes (random), i.e., 43% of the initial nodes in the field.

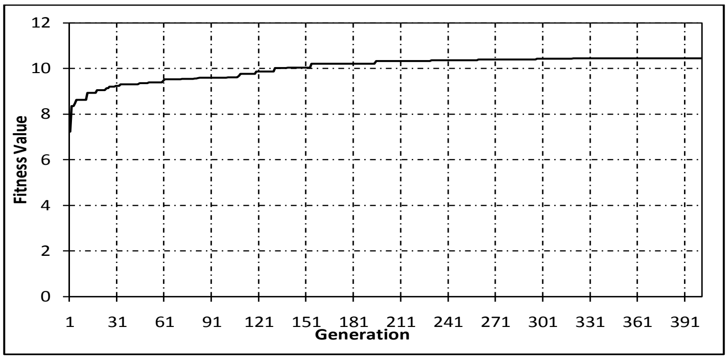


FIGURE 7.4: Best Fitness Value Vs Generation

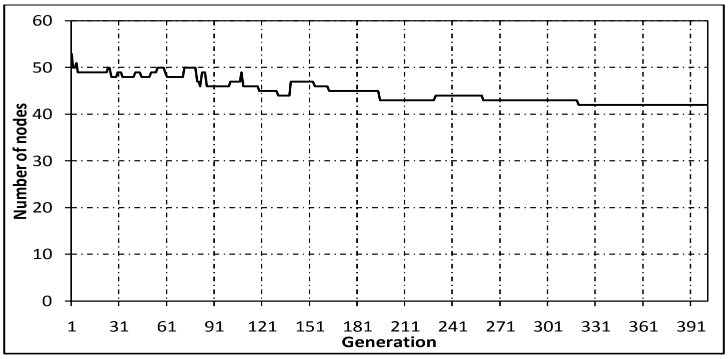


FIGURE 7.5: Active Nodes Vs Generation

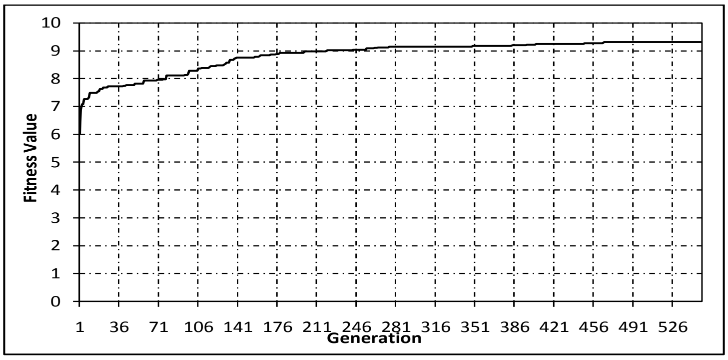


FIGURE 7.6: Best Fitness Value Vs Generation

7.4.2 For Minimum 20% Overlapped Area

a. For 50×50 m^2 Area with 25 Nodes: Figure 7.8 shows the best fitness value over the generation. There is an increase in the best fitness value of the current population over generations. After 140 generations the best fitness value is constant, .i.e. genetic

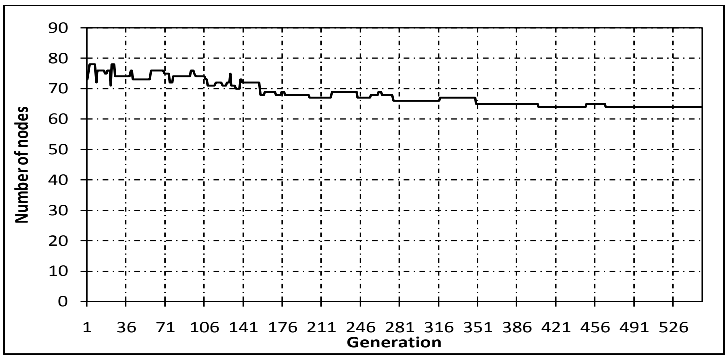


Figure 7.7: Active Nodes Vs Generation

algorithm applied optimizes the problem. As the fitness function is increasing the value of the best fitness value so the number of active nodes is decreasing to an optimal value as shown in figure 7.9. The results show that the solution for the stated problem converges at 13 nodes (random), i.e. 52% of the initial nodes in the field.

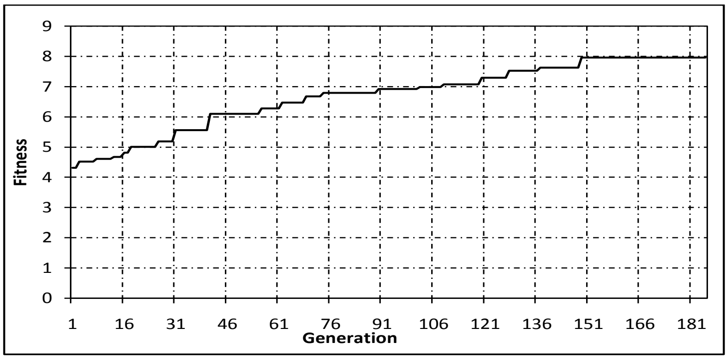


Figure 7.8: Best Fitness Value Vs Generation

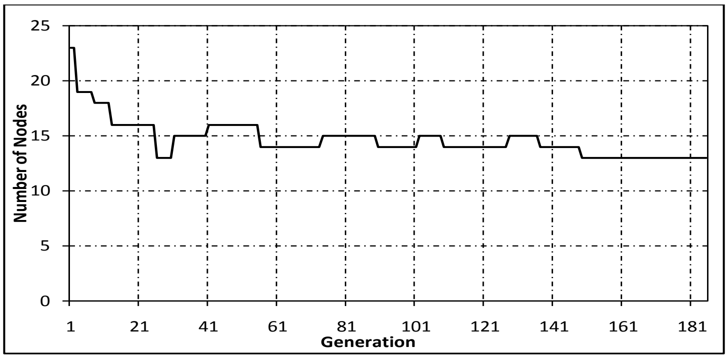


Figure 7.9: Active Nodes Vs Generation

b. For 100×100 m^2 with 100 Nodes: Figure 7.10 shows an increase in the best fitness value of the population over generations. Genetic algorithm has a constant best fitness value after 340 generations. Figure 7.11 shows that the number of active nodes is decreasing with the generations and the solution for the stated problem is converging at 45 nodes (random), i.e. 45% of the initial nodes in the field.

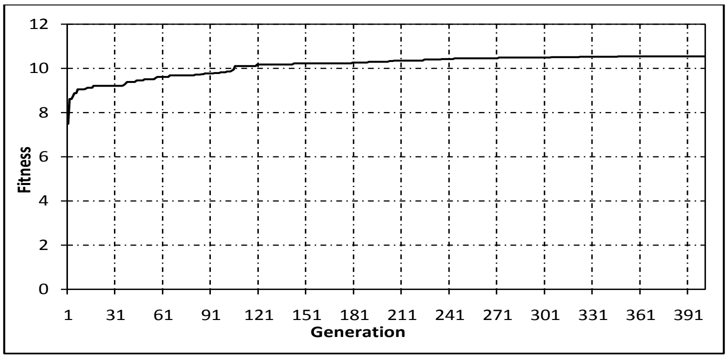


Figure 7.10: Best Fitness Value Vs Generation

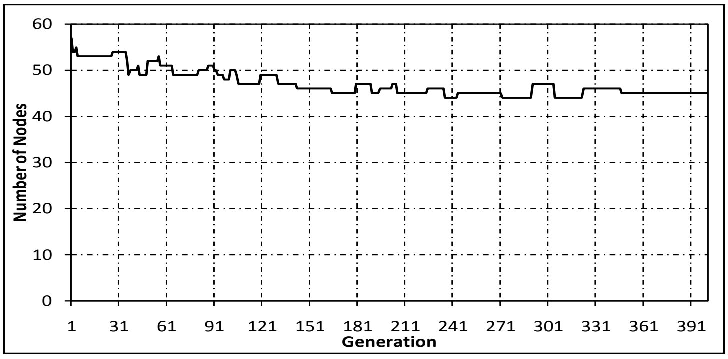


Figure 7.11: Active Nodes Vs Generation

c. For 125×125 m^2 Area with 150 Nodes: Figure 7.12 shows the increase in the best fitness value over generations. After 300 generations the increase is slight and after 450 generations it converges. The results show that the solution for the stated problem converges at 70 nodes (random), i.e., 47% of the initial nodes in the field, as shown in figure 7.13.

Table 7.2 shows that 48% to 52% of the initial nodes (random) are required to complete the requirements of sensing the whole area with minimum overlapping in the sensed area of 20% to 30% respectively.

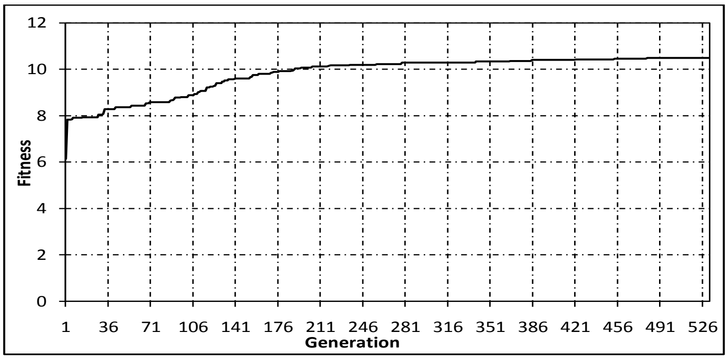


Figure 7.12: Best Fitness Value Vs Generation

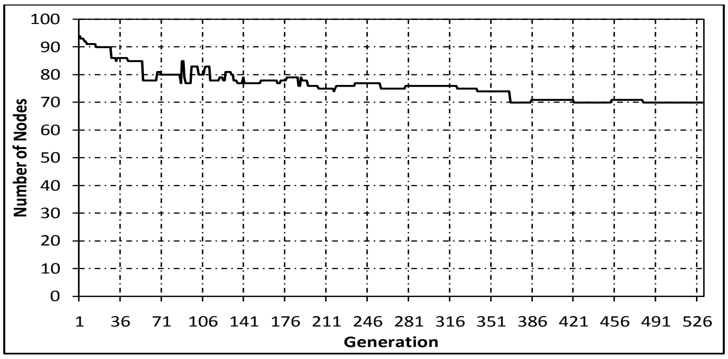


Figure 7.13: Active Nodes Vs Generation

Table 7.2: Result Summary

Nodes	**Area**	**Minimum No. of Nodes for 20% Overlapped area**		**Minimum No. of Nodes for 30% Overlapped area**	
		From (7.1)	**Random**	**From (7.1)**	**Random**
25	50 X 50	10	12	11	13
100	100 X 100	39	42	41	45
150	125 X 125	60	64	65	70

7.4.3 Application

The sensor nodes send their information about the location to base station once. Size of the data is very small, so it does not consume much energy. Hence the energy efficiency issue of the applied approach is not affected. Base station performs the optimization. The network knows in advance the minimum number of active nodes that will complete the requirements of the application.

In sleep/wake-up scheduling [12] approaches, if the selected active nodes are less than the calculated minimum active nodes, the result produced by active nodes will be unreliable. Scheduling approaches should ensure that the number of selected nodes for a particular time slice should be higher than the acceptance level. The reliability of clustering algorithms [13, 14] highly depends on the covered area and the redundancy in data. The network should be left with acceptable number of active nodes otherwise it should inform the base station.

7.5 Conclusion

In a wireless sensor network, the reliability of a network heavily depends on the sensed area and the redundancy in data. Due to the dynamic nature of the wireless sensor networks nodes die randomly in the field. Node death in field causes decrease in redundancy of data as well as reduction in the sensed area. The work of this chapter finds when network transit from reliable state to unreliable state over the time. Work of chapter optimizes the problem to find the minimum number of active nodes that will sense almost complete area with the minimum acceptable redundancy in the data. Multi Objective Genetic Algorithm based approach is applied to find the state of the network. Number of alive nodes, total sensed area and total overlapped area are the parameters for fitness function of proposed GA approach.

Chapter 8

Conclusion and Future Aspects

Wireless sensor networks have gained ample interest because of their wide range of applications. Efficient energy consumption of nodes to prolong the network lifetime is the principal design aspect for wireless sensor networks. This chapter highlights the contribution of our work along with the opportunities for future work.

8.1 Summary of Contribution

Wireless sensor networks have recently gained the attention of researchers. WSNs have unique characteristics that make them different from other types of the networks. Sensor nodes are energy constrained so that energy conservation of the nodes is one of the most important characteristics of WSNs. Clustering algorithms is a popular hierarchical architecture that conserve energy of nodes along with fault-tolerance and scalability. Clustering approach groups nodes in different clusters and selects at least one node as cluster head for each cluster. Cluster heads gather and aggregates the data from member nodes. Then cluster heads transmit the aggregated data to base station directly or multi-hop.

In this thesis, the main motivation was to prolong the network lifetime with load balanced network with clustering approach as base. Various issues regarding the performance of clustering algorithms were explored and various solutions to prolong the network lifetime have been proposed and evaluated for various wireless sensor network scenarios. Chapter 4 exploited the fix round-time problem and proposed a network adaptive variable round-time solution. Further, to load balance the network a new cluster head selection solution was proposed that chose cluster head according to their location in the application area. chapter 5 exploited the network heterogeneity caused by either network dynamics or initially few nodes have more energy than others. Proposed solution gave more preference to high remaining energy nodes than low energy nodes for cluster head selection. Chapter 6 gave due importance to balanced size clusters. The proposed solution formed balanced clusters depending on the two thresholds. Chapter 7 addressed the reliability state of network. Next sections have summarized the main contributions and achievements of the proposed solutions.

8.1.1 Clustering Algorithm to Load Balance the WSNs

A load balanced clustering approach was presented for data centric wireless sensor networks that had network adaptive round-time and a cluster head selection method to reduce intra-cluster communication distance. Clustering algorithms operate in rounds to load balance the network by rotating the role of cluster head among the nodes. Classical clustering algorithms had fix round-time value that has been calculated from initial parameters of networks. Work of this chapter presented a dynamic round-time clustering solution that determined round-time according to number of alive nodes. The proposed solution had very small round-time for last few nodes that consumed more energy from these nodes because of frequent re-clustering, consequently diminished the performance of network. An enhancement over proposed dynamic round-time clustering solution was presented named network adaptive round-time solution.

Proposed network adaptive round-time solution had round-time depending upon the number of alive nodes. To take care of low remaining energy nodes of last rounds proposed network adaptive round-time solution had fix round-time after the death of 50% nodes. Consequently, proposed solution had taken into account the tradeoffs of long and short round-time and round-time was evaluated dynamically to adapt the network dynamics. We compared our proposed network adaptive round-time solution to LEACH, which had fixed round-time throughout the network lifetime and another proposed dynamic round-time solution, which had dynamic network lifetime throughout the network lifetime. Three different topologies of different sizes were simulated to analyze proposed network adaptive round-time solution and to compare with LEACH and proposed dynamic round-time solution. Network lifetime and data units received at base station metrics were reviewed. There was an improvement of 7% for both network lifetime and data units received at base station in case of proposed network adaptive round-time solution over LEACH. The proposed network adaptive round-time scheme outperformed both the remaining approached in network lifetime and data gathered at base station.

Proposed solution of cluster head selection divided the network into two parts: Border area and Inner area. Only inner area nodes could contend for the role of cluster head and hence border area nodes were never selected as a cluster head. Proposed solution had cluster heads near to center of cluster hence as a result of that had less intra-cluster communication distance. Partitioning distance was quantify by simulating the network for various different values. Node death rate, energy consumption rate, network lifetime and data units received at base station metrics were reviewed to analyze the performance of proposed solution and to compare with LEACH. Proposed solution had lower node death and energy consumption rate than LEACH for all the simulated topologies. There was an increase of 32% for first node death and 8% for data units received at base station for one of topology. Simulation results had ensured that the proposed solution had better load balanced network and longer network lifetime than that of LEACH with more data units received successfully at base station.

8.1.2 Capitalizing Heterogeneity of Sensor Nodes to Prolong The Network Lifetime

Node heterogeneity could happen either by deploying few resource-rich nodes or by network dynamics. Heterogeneity of nodes should be capitalized to protract network lifetime. We presented a cluster head solution that selected cluster heads with consideration of both residual node-energy and average network-energy. The main goal of proposed solution was to select high residual energy nodes in the early rounds of an epoch as well as for the current round. We examined the effectiveness of proposed cluster head selection with LEACH and SEP protocols for cluster head selection. Evaluation affirmed that the proposed solution had selected nodes with high remaining energy as the role of cluster head for the current round-time as compared to other considered approaches.

Proposed solution had load balanced network by electing high residual energy-nodes as cluster heads in the early rounds of an epoch and left the remaining nodes (nodes that are not selected as cluster head due to lower residual node-energy) as the role of cluster head for the last rounds of the epoch. Proposed solution contended with the both type of node heterogeneity.

Both homogeneous and heterogeneous sensor network topologies were generated and were used to examine the effectiveness of proposed solution via simulation experiments. Three different network topologies of different sizes and number of nodes were simulated for homogeneous network. Network lifetime and data units received at base station metrics were selected to analyze the performance of proposed solution. There was increase of 10%, 12% and 14% for first node death, half node death and last node death time respectively in case of proposed solution over LEACH. Simulation results also affirmed that proposed solution also outperform SEP also completely. Network topologies varying number of heterogeneous nodes and initial energy were simulated for proposed solution, LEACH and SEP. For all network topologies, network lifetime and data units received at base station had significant improvement in case of proposed solution over LEACH and SEP. Proposed solution capitalized the node heterogeneity for cluster head selection and protracted the network lifetime.

8.1.3 Balanced-Size Clusters to Extend Network Lifetime

We presented a cluster formation algorithm to address the unbalanced cluster size formation problem. Clustering algorithms operated in round with each round having two phases namely set-up phase and data transmission phase. Data transmission phase had been divided into frames and frames were further segregated in time slots. Random cluster head selection headed the network to unequal size clusters. Different amount of energy were consumed by clusters of uneven sizes that made the network load unbalanced. Small sized clusters consumed more energy as compared to large sized clusters. Frame length of small size clusters has been small than large size clusters therefore time slots to transmit data per node in the data transmission phase was more for small size clusters.

Proposed solution divided the cluster formation process into two phases, initial phase and rescue phase, to set up size-balanced clusters. Cluster formation was completed according to two threshold, $Th_{cluster}$ for number of nodes in clusters and $Th_{distance}$ for joining the cluster head. In the initial phase, cluster heads had restriction on number of member nodes ($Th_{cluster}$) and a cluster could have number of nodes not more than $Th_{cluster}$, consequently few nodes were left without any cluster head. An algorithm was proposed for rescue phase which made selection of cluster head nodes for un-clustered nodes by capitalizing the $Th_{cluster}$, $Th_{distance}$ and JOIN condition. An un-clustered node has been selecting a cluster head if it was not located far than $Th_{distance}$ and had member nodes less than $Th_{cluster}$. If the JOIN condition was not satisfied by ant cluster head then nearest cluster head has been selected as cluster head.

Cluster quality in term of number of member nodes in each cluster and total communication distance of clusters was examined to find the value of $Th_{distance}$ for proposed solution. Analysis affirmed that the proposed solution had better cluster quality than the conventional clustering algorithm. Proposed balanced size clusters solution was compared with LEACH. Network Lifetime and data units received at base station metrics were chosen to analyze the performance. Simulation results asserted that proposed algorithm had clusters of balanced size and had prolonged network lifetime.

8.1.4 Reliability of Wireless Sensor Networks

A wireless sensor network needed, for successful operation, that the deployed area should be thoroughly sensed by the all nodes collectively, i.e., there should not be any sensing hole in the network. There should be spatial redundancy in the sensed data to obtain quality data and the dense deployment of nodes has been providing the much needed spatial redundancy. Reliability of wireless sensor networks was measured in terms of covered area by all active nodes and spatial redundancy in sensed data. We presented a genetic algorithm based approach that analyzed the reliability state of wireless sensor networks. The main aim of the proposed solution was to select minimum number of nodes (random) that sense almost all area with the desired redundancy in sensed data. We examined the productiveness of proposed solution by consideration of various network topologies. Three different network topologies varying in size and number of nodes were simulated. The results asserted the convergence of proposed solution. The outcome of presented solution has been helpful to find when a reliable state network will transit to unreliable state.

8.2 Future Work

A lot can be achieved by extending the current studies to energy efficient wireless sensor networks. Few of the future work directions are:

- As discussed in the presented work, position of cluster head is decisive for the performance of clustering algorithms and heterogeneous nodes are gaining considerable importance. Position of heterogeneous nodes in the network is a vital issue for heterogeneous clustering algorithms efficiency. Heterogeneous nodes are selected more times than the other nodes so the location of these nodes in the field is decisive for heterogeneous clustering algorithms. The area of positioning of heterogeneous nodes is wide open.
- We constructed size-balanced clusters by implementing $\text{Th}_{cluster}$ and $\text{Th}_{distance}$. Other parameters, such as RSS, for threshold can play an important role to achieve balanced clusters while reducing overhead. The proposed solution presented in the thesis is a fully distributed approach while a centralized approach for equal sized cluster is yet to explore.
- One of the underlined assumptions for the accomplished research work is that the sensors in the deployed area and the base station are stationary but there is also an increasing interest, in the research community, about the mobile sensors and base stations. And to inspect the reliability of such mobile network is a critical issue.
- Clustering based algorithms primarily focus on cluster head selection, cluster formation and data transmission from member nodes to cluster heads and cluster heads to base station, but give no importance to secure communication. Secure communication of data is required for clustering based algorithms and gaining a rapid interest among research community.
- Wireless sensor networks generate enormous amount of data during the complete lifetime. Mining of that provides useful information and helps to analyze the scenario more interestingly. Mining of information can also help to reduce energy consumption. Outlier detections are gaining ample interest for energy conservation along with security issue of network of nodes and are yet to be explored in detail.

Bibliography

[1] C.-Y. Chong and S. P. Kumar, "Sensor Networks: Evolution, Opportunities, and Challenges," *Proceedings of the IEEE*, vol. 91, no. 8, pp. 1247 – 1256, Aug. 2003.

[2] Q. Wang and I. Balasingham, *Wireless Sensor Networks - An Introduction, Wireless Sensor Networks: Application - Centric Design.* Yen Kheng Tan (Ed.), 2010.

[3] U. B. Desai and I. Balasingham, "Wireless Sensor Networks: Technology Roadmap."

[4] "The Evolution of Wireless Sensor Networks," pp. 1 – 5.

[5] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," *Computer Networks*, vol. 38, no. 4, pp. 393 – 422, 2002.

[6] G. Anastasi, M. Conti, M. Di Francesco, and A. Passarella, "Energy conservation in wireless sensor networks: A survey," *Ad Hoc Netw.*, vol. 7, no. 3, pp. 537–568, May 2009.

[7] H. Karl and A. Willig, *Protocols and architectures for wireless sensor networks.* John Wiley & Sons, 2007.

[8] J. Zheng and A. Jamalipour, *Wireless Sensor Networks A Networking Perspective.* John Wiley & Sons, 2009.

[9] Y. Li, M. T. Thai, and W. Wu, *Wireless Sensor Networks and Applications.* Springer, 2008.

[10] E. ayrc and C. Rong, *Security in Wireless Ad Hoc and Sensor Networks Security.* John Wiley & Sons, 2009.

[11] J. Carle and D. Simplot-Ryl, “Energy-efficient area monitoring for sensor networks,” *Computer*, vol. 37, pp. 40 – 46, 2004.

[12] A. Milenković, C. Otto, and E. Jovanov, “Wireless sensor networks for personal health monitoring: Issues and an implementation,” *Comput. Commun.*, vol. 29, no. 13-14, pp. 2521–2533, Aug. 2006.

[13] L. Selavo, A. D. Wood, Q. Cao, T. I. Sookoor, H. Liu, A. Srinivasan, Y. Wu, W. Kang, J. A. Stankovic, D. Young, and J. Porter, “LUSTER: wireless sensor network for environmental research,” in *SenSys*, 2007, pp. 103–116.

[14] A. Mainwaring, D. Culler, J. Polastre, R. Szewczyk, and J. Anderson, “Wireless sensor networks for habitat monitoring,” in *Proceedings of the 1st ACM international workshop on Wireless sensor networks and applications*, ser. WSNA ’02, 2002, pp. 88–97.

[15] C. Buragohain, D. Agrawal, and S. Suri, “Distributed navigation algorithm for sensor networks,” in *In Proceedings of the 25th IEEE Conference on Computer Communication (INOCOM’06)*, 2006, pp. 103–116.

[16] J. Tavares, F. J. Velez, and J. Ferro, “Application of wireless sensor networks to automobiles,” *Measurement Science Review*, vol. 8, no. 3, pp. 65–70, June 2008.

[17] A. Flammini, P. Ferrari, D. Marioli, E. Sisinni, and A. Taroni, “Wired and wireless sensor networks for industrial applications,” *Microelectron. J.*, vol. 40, no. 9, pp. 1322–1336, Sep. 2009.

[18] A. Arora, P. Dutta, S. Bapat, V. Kulathumani, H. Zhang, V. Naik, V. Mittal, H. Cao, M. Demirbas, M. Gouda, Y. Choi, T. Herman, S. Kulkarni, U. Arumugam, M. Nesterenko, A. Vora, and M. Miyashita, “A line in the sand: a wireless sensor network for target detection, classification, and tracking,” *Comput. Netw.*, vol. 46, no. 5, pp. 605–634, Dec. 2004.

[19] R. Roman, C. Alcaraz, and J. Lopez, “The role of wireless sensor networks in the area of critical information infrastructure protection,” *Inf. Secur. Tech. Rep.*, vol. 12, no. 1, pp. 24–31, Mar. 2007.

[20] A. A. Abbasi and M. Younis, “A survey on clustering algorithms for wireless sensor networks,” *Comput. Commun.*, vol. 30, no. 14-15, pp. 2826–2841, Oct. 2007.

[21] D. Wei and H. Chan, “Clustering ad hoc networks: Schemes and classifications,” *3rd Annual IEEE Communications Society on Sensor and Ad Hoc Communications and Networks*, vol. 3, pp. 920–926, 2006.

[22] O. Younis, M. Krunz, and S. Ramasubramanian, “Node clustering in wireless sensor networks: Recent developments and deployment challenges,” *IEEE Network Magazine*, vol. 20, pp. 20–25, 2006.

[23] A. Hac, *Wireless Sensor Network Designs.* John Wiley & Sons, 2003.

[24] M. M. Afsar and M.-H. Tayarani-N, “Clustering in sensor networks: A literature survey,” *Journal of Network and Computer Applications*, vol. 46, pp. 198 – 226, 2014.

[25] R. Rajagopalan and P. K. Varshney, “Data aggregation techniques in sensor networks: A survey,” *Comm. Surveys and Tutorials, IEEE*, vol. 8, pp. 48–63, 2006.

[26] M. Yarvis, A. Kushalnagar, H. Singh, Y. Liu, and S. Singh, “Exploiting heterogeneity in sensor networks,” in *in Proc. of the IEEE Infocom*, 2005.

[27] J. Kabara and M. Calle, “MAC Protocols Used by Wireless Sensor Networks and a General Method of Performance Evaluation,” *International Journal of Distributed Sensor Networks*, vol. 2012, pp. 1 – 11, 2012.

[28] N. A. Pantazis and D. D. Vergados, “A survey on power control issues in wireless sensor networks.” *IEEE Communications Surveys and Tutorials*, vol. 9, no. 1-4, pp. 86–107, 2007.

[29] D. Corbett, A. Ruzzelli, and U. of Sydney. School of Information Technologies, *A Procedure for Benchmarking Mac Protocols Used in Wireless Sensor Networks: Technical Report No. 593*, ser. Technical report (University of Sydney. School of Information Technologies). University of Sydney - School of Information Technologies, 2006.

[30] I. Demirkol, C. Ersoy, and F. Alagoz, “MAC protocols for wireless sensor networks: a survey,” *IEEE Communications Magazine*, vol. 44, no. 4, pp. 115–121, Apr. 2006.

[31] A. El-Hoiydi, J.-D. Decotignie, C. Enz, and E. Le Roux, "Poster Abstract: WiseMAC, an Ultra Low Power MAC Protocol for the wiseNET Wireless Sensor Network," in *Proceedings of the 1st International Conference on Embedded Networked Sensor Systems*, ser. SenSys '03, 2003, pp. 302–303.

[32] A. El-Hoiydi and J.-D. Decotignie, "WiseMAC: An Ultra Low Power MAC Protocol for Multi-hop Wireless Sensor Networks." in *ALGOSENSORS*, ser. Lecture Notes in Computer Science, vol. 3121. Springer, 2004, pp. 18–31.

[33] A. El-Hoiydi, "Spatial TDMA and CSMA with preamble sampling for low power ad hoc wireless sensor networks." in *ISCC*. IEEE Computer Society, 2002, pp. 685–692.

[34] W. Ye, J. Heidemann, and D. Estrin, "An Energy-Efficient MAC Protocol for Wireless Sensor Networks," in *IEEE INFOCOM 2002*, 2002, pp. 1567 – 1576.

[35] V. Bharghavan, A. Demers, S. Shenker, and L. Zhang, "MACAW: A Media Access Protocol for Wireless LAN's," in *Proceedings of the Conference on Communications Architectures, Protocols and Applications*, ser. SIGCOMM '94, 1994, pp. 212–225.

[36] V. Rajendran, K. Obraczka, and J. J. Garcia-Luna-Aceves, "Energy-efficient, collision-free medium access control for wireless sensor networks," *Wirel. Netw.*, vol. 12, no. 1, pp. 63–78, Feb. 2006.

[37] W. L. Lee, A. Datta, and R. Cardell-Oliver, "Flexitp: A flexible-schedule-based tdma protocol for fault-tolerant and energy-efficient wireless sensor networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 19, no. 6, pp. 851–864, 2008.

[38] A. Boukerche, X. Cheng, and J. Linus, "A performance evaluation of a novel energy-aware data-centric routing algorithm in wireless sensor networks." *Wireless Networks*, vol. 11, no. 5, pp. 619–635, 2005.

[39] M. Maroti, B. Kusy, G. Simon, and k. Ldeczi, "Robust multi-hop time synchronization in sensor networks." in *International Conference on Wireless Networks*. CSREA Press, 2004, pp. 454–460.

[40] I. Rhee, A. Warrier, M. Aia, and J. Min, "Z-MAC: A Hybrid MAC for Wireless

Sensor Networks," in *Proceedings of the 3rd International Conference on Embedded Networked Sensor Systems*, ser. SenSys '05, 2005, pp. 90–101.

[41] I. Rhee, A. C. Warrier, and L. Xu, "Randomized dining philosophers to tdma scheduling in wireless sensor networks," Tech. Rep., 2004.

[42] D. Tian and N. D. Georganas, "A node scheduling scheme for energy conservation in large wireless sensor networks," in *Wireless Networks and Mobile Computing*, 2003, pp. 1 – 37.

[43] Y. Xu, J. Heidemann, and D. Estrin, "Geography-informed energy conservation for ad hoc routing," in *Proceedings of the 7th Annual International Conference on Mobile Computing and Networking*, ser. MobiCom '01, 2001, pp. 70–84.

[44] C. Schurgers, V. Tsiatsis, and M. B. Srivastava, "STEM: Topology Management for Energy Efficient Sensor Networks," in *IN IEEE AEROSPACE CONFERENCE*, 2002, pp. 78–89.

[45] A. Keshavarzian, H. Lee, and L. Venkatraman, "Wakeup scheduling in wireless sensor networks." in *MobiHoc*. ACM, 2006, pp. 322–333.

[46] N. Tezcan and W. Wang, "TTS: A Two-Tiered Scheduling Algorithm for Effective Energy Conservation in Wireless Sensor Networks," in *ICC*, 2006, pp. 3359–3364.

[47] A. Cerpa and D. Estrin, "ASCENT: Adaptive Self-Configuring sEnsor Networks Topologies." *IEEE Trans. Mob. Comput.*, vol. 3, no. 3, pp. 272–285, 2004.

[48] B. Chen, K. Jamieson, H. Balakrishnan, and R. Morris, "Span: An energy-efficient coordination algorithm for topology maintenance in ad hoc wireless networks," *Wirel. Netw.*, vol. 8, no. 5, pp. 481–494, Sep. 2002.

[49] M. Cardei, M. T. Thai, Y. Li, and W. Wu, "Energy-efficient target coverage in wireless sensor networks," in *in IEEE Infocom*, 2005, pp. 1976–1984.

[50] D. Tian and N. D. Georganas, "A coverage-preserving node scheduling scheme for large wireless sensor networks." in *WSNA*. ACM, 2002, pp. 32–41.

[51] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proceedings of the 33rd*

Hawaii International Conference on System Sciences-Volume 8, ser. HICSS '00. Washington, DC, USA: IEEE Computer Society, 2000, pp. 10 – 20.

[52] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *Wireless Communications, IEEE Transactions on*, vol. 1, no. 4, pp. 660 – 670, oct 2002.

[53] T. Murata and H. Ishibuchi, "Performance evaluation of genetic algorithms for flowshop scheduling problems," in *Proceedings of the First IEEE Conference on IEEE World Congress on Computational Intelligence, Evolutionary Computation, 1994.*, jun 1994, pp. 812 – 817.

[54] W. B. Heinzelman, "Application-specific protocol architectures for wireless networks," Ph.D. dissertation, 2000, aAI0801929.

[55] K.-Y. Jang, K. T. Kim, and H. Y. Youn, "An energy efficient routing scheme for wireless sensor networks." in *ICCSA Workshops.* IEEE Computer Society, 2007, pp. 399 – 404.

[56] C. Jing, T. GU, and L. Chang, "ESCAL: An Energy-Saving Clustering Algorithm Based on LEACH," in *IEEE Int. Symposium on Knowledge Acquisition and Modeling Workshop*, 2008, pp. 403 – 406.

[57] M. Bsoul, A. Al-Khasawneh, A. E. Abdallah, E. E. Abdallah, and I. Obeidat, "An energy-efficient threshold-based clustering protocol for wireless sensor networks," *Wirel. Pers. Commun.*, vol. 70, no. 1, pp. 99–112, May 2013.

[58] F. Bajaber and I. Awan, "Adaptive decentralized re-clustering protocol for wireless sensor networks," *Journal of Computer System and Science*, vol. 77, no. 2, pp. 282–292, Mar. 2011.

[59] ——, "Energy efficient clustering protocol to enhance lifetime of wireless sensor network," *Journal of Ambient Intelligence and Humanized Computing*, vol. 1, no. 4, pp. 239–248, 2010.

[60] A. Azim and M. Islam, "A dynamic round-time based fixed low energy adaptive clustering hierarchy for wireless sensor networks," in *IEEE 9th Malaysia International Conference on Communications (MICC), 2009*, 2009, pp. 922–926.

[61] Azim and M. Islam, "Dynamic service policy-based clustered wireless sensor networks," in *IEEE 6th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), 2010*, 2010, pp. 196–202.

[62] G. Chen, C. Li, M. Ye, and JieWu, "An unequal cluster-based routing protocol in wireless sensor networks," *Wireless Networks*, vol. 15, no. 2, pp. 193 – 207, 2009.

[63] J. Chang-Jiang, S. Wei-ren, X. min, and T. Xian-lun, "Energy-balanced unequal clustering protocol for wireless sensor networks," *The Journal of China Universities of Posts and Telecommunications*, vol. 17, no. 4, pp. 94–99, 2010.

[64] J. Tillett, R. Rao, and F. Sahin, "Cluster-head identification in ad-hoc sensor netwoks using particle swarm optimization," in *International Conference on Personal Wireless Communications (ICPWC'02)*, New Delhi, India, DEC 2002.

[65] S. Ghiasi, A. Srivastava, X. Yang, and M. Sarrafzadeh, "Optimal energy aware clustering in sensor networks," *Sensors*, vol. 2, pp. 258–269, 2002.

[66] S. Hussain and A. W. Matin, "Base station assisted hierarchical cluster-based routing." in *ICWMC*. IEEE Computer Society, 2006.

[67] S. Bandyopadhyay and E. J. Coyle, "An energy efficient hierarchical clustering algorithm for wireless sensor networks," 2003.

[68] G. Pei and C. Chien, "Low power TDMA in large wireless sensor networks," in *IEEE Military Comm. Conf. (MILCOM 01)*, vol. 1, 2001.

[69] M. Gerla, T. Kwon, and G. Pei, "On demand routing in large ad hoc wireless networks with passive clustering," 2000.

[70] A. S. Zahmati, B. Abolhassani, A. Asghar, B. Shirazi, and A. S. Bakhtiari, "An energy-efficient protocol with static clustering for wireless sensor networks," *World academy of Science, Engineering and Technology*, 2007.

[71] A. Durresi, V. Paruchuri, and L. Barolli, "Clustering protocol for sensor networks," in *AINA06.*, 2006.

[72] S.-G. Deng, L.-F. Shen, and X.-R. Zhu, "Energy-efficient data aggregation protocol based on static clustering for wireless sensor networks," in *PIERS Proceedings*,

2008, pp. 470 – 473.

[73] W. Chen, W. Li, H. Shou, and B. Yuan, "A QoS-based adaptive clustering algorithm for wireless sensor networks," in *IEEE International Conference on Mechatronics and Automation*, 2006, pp. 1947 – 1952.

[74] A. S. Raghuvanshi, S. Tiwari, R. Tripathi, and N. Kishor, "Optimal number of clusters in wireless sensor networks: a FCM approach." *International Journal of Sensor Networks*, vol. 12, no. 1, pp. 16–24, 2012.

[75] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms.* Norwell, MA, USA: Kluwer Academic Publishers, 1981.

[76] D. E. Goldberg, *Genetic Algorithms in Search and Machine Learning.* Pearson Education, 2006.

[77] M. Gen and R. Cheng, *Genetic Algorithms and Engineering Optimization (Engineering Design and Automation).* Wiley-Interscience, 1999.

[78] J.-L. Liu and C. V. Ravishanka, "LEACH-GA: Genetic Algorithm-Based Energy-Efficient Adaptive Clustering Protocol for Wireless Sensor Networks," *International Journal of Machine Learning and Computing*, vol. 1, no. 1, pp. 79 – 85, December 2011.

[79] S. Hussain, A. W. Matin, and O. Islam, "Genetic algorithm for energy efficient clusters in wireless sensor networks." in *ITNG.* IEEE Computer Society, 2007, pp. 147–154.

[80] ——, "Genetic algorithm for hierarchical wireless sensor networks." *JNW*, vol. 2, no. 5, pp. 87–97, 2007.

[81] E. Heidari and A. Movaghar, "An efficient method based on genetic algorithms to solve sensor network optimization problem," *International journal on application of graph theory in wireless ad hoc networks and sensor networks*, 2011.

[82] V. Godbole, "Genetic algorithm-based energy-efficient protocol for wireless sensor networks," *International journal of ultra wideband communications and systems : IJUWBCS*, vol. 2, no. 3, pp. 133–140, December 2012.

[83] G. Smaragdakis, I. Matta, and A. Bestavros, “SEP: A Stable Election Protocol for clustered heterogeneous wireless sensor networks,” in *Second International Workshop on Sensor and Actor Network Protocols and Applications (SANPA 2004)*, Boston, MA, August 2004.

[84] O. Rehman, N. Javaid, B. Manzoor, A. Hafeez, A. Iqbal, and M. Ishfag, “Energy Consumption Rate based Stable Election Protocol (ECRSEP) for WSNs,” in *Body Area Sensor Networks (BASNet), 2013 International Workshop on*, 2013, pp. 932–937.

[85] F. Aderohunmu and J. Deng, “An Enhanced Stable Election Protocol (SEP) for Clustered Heterogeneous WSN,” in *ISDCA*, 2006.

[86] L. Qing, Q. Zhu, and M. Wang, “Design of a distributed energy-efficient clustering algorithm for heterogeneous wireless sensor networks,” *Comput. Commun.*, vol. 29, no. 12, pp. 2230–2237, Aug. 2006.

[87] D. Kumar, T. C. Aseri, and R. Patel, “EEHC: Energy efficient heterogeneous clustered scheme for wireless sensor networks,” *Computer Communications*, vol. 32, no. 4, pp. 662 – 667, 2009.

[88] D. Kumar, R. Patel, and T. C. Aseri, “Prolonging network lifetime and data accumulation in heterogeneous sensor networks.” *Int. Arab J. Inf. Technol.*, vol. 7, no. 3, pp. 302–309, 2010.

[89] O. Rehman, N. Javaid, B. Manzoor, A. Hafeez, A. Iqbal, and M. Ishfag, “ED-DEEC: Enhanced Developed Distributed Energy-Efficient Clustering for Heterogeneous Wireless Sensor Networks,” in *Bosy Area Sensor Networks (BASNet), 2013 International Workshop on*, 2013, pp. 914–919.

[90] R. Tandon, B. Dey, and S. Nandi, “Weight based clustering in wireless sensor networks,” in *Communications (NCC), 2013 National Conference on*, Feb 2013, pp. 1–5.

[91] R. Tandon and S. Nandi, “CRP: Cluster head Re-election Protocol for heterogeneous wireless sensor networks,” in *Communication Systems and Networks (COMSNETS), 2013 Fifth International Conference on*, Jan 2013, pp. 1–10.

[92] D. Kumar, "Performance analysis of energy efficient clustering protocols for maximising lifetime of wireless sensor networks," *Wireless Sensor Systems, IET*, vol. 4, no. 1, pp. 9–16, March 2014.

[93] M. Liu, J. Cao, G. Chen, and X. Wang, "An energy-aware routing protocol in wireless sensor networks," *Sensors*, vol. 9, no. 1, pp. 445–462, 2009.

[94] E. Duarte-Melo and M. Liu, "Analysis of energy consumption and lifetime of heterogeneous wireless sensor networks," in *Global Telecommunications Conference, 2002. GLOBECOM '02. IEEE*, vol. 1, Nov 2002, pp. 21–25 vol.1.

[95] "An efficient multi-level clustering approach for a heterogeneous wireless sensor network using link correlation," *EURASIP Journal on Wireless Communications and Networking*, vol. 2014, no. 1, pp. 1–10, 2014.

[96] B. A. Attea and E. A. Khalil, "A new evolutionary based routing protocol for clustered heterogeneous wireless sensor networks," *Applied Soft Computing*, vol. 12, no. 7, pp. 1950 – 1957, 2012.

[97] T. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. Upper Saddle River, NJ, USA: Prentice Hall PTR, 2001.

[98] S. Halder, A. Ghosal, and S. D. Bit, "A pre-determined node deployment strategy to prolong network lifetime in wireless sensor network," *Computer Communications*, vol. 34, no. 11, pp. 1294 – 1306, 2011.

[99] C. SHA and R. chuan WANG, "Energy-efficient node deployment strategy for wireless sensor networks," *The Journal of China Universities of Posts and Telecommunications*, vol. 20, no. 1, pp. 54 – 57, 2013.

[100] J. Vales-Alonso, F. Parrado-Garca, P. Lpez-Matencio, J. Alcaraz, and F. Gonzlez-Castao, "On the optimal random deployment of wireless sensor networks in non-homogeneous scenarios," *Ad Hoc Networks*, vol. 11, no. 3, pp. 846 – 860, 2013.

[101] S. Mccanne, S. Floyd, and K. Fall, "ns2 (network simulator 2)." [Online]. Available: http://www.isi.edu/nsnam/ns/

[102] "Nctuns: Network simulator and emulator." [Online]. Available: http://nsl.csie.nctu.edu.tw/nctuns.html.

[103] "Mannasim." [Online]. Available: http://www.mannasim.dcc.ufmg.br/.

[104] "The omnet++ discrete event simulator system." [Online]. Available: http://www.omnetpp.org/

[105] "Opnet tech inc." [Online]. Available: http://www.opnet.com/

[106] "Glomosim: Global mobile information systems simulation library." [Online]. Available: http://pcl.cs.ucla.edu/projects/glomosim/.

[107] T. Issariyakul and E. Hossain, *Introduction to Network Simulator NS2*, 1st ed. Springer Publishing Company, Incorporated, 2008.

[108] "μamps extension ns code extensions for leach." [Online]. Available: http://www-mtl.mit.edu/researchgroups/icsystems/uamps/research/leach/leach_code.shtml

[109] R. Pratap, *Getting Started with MATLAB: A Quick Introduction for Scientists and Engineers*, 1st ed. New York, NY, USA: Oxford University Press, Inc., 2002.

[110] G. Han, H. Xu, T. Q. Duong, J. Jiang, and T. Hara, "Localization algorithms of wireless sensor networks: a survey," *Telecommunication Systems*, vol. 52, no. 4, pp. 2419 – 2436, April 2013.

[111] L. Wang and Q. Xu, "Gps-free localization algorithm for wireless sensor networks," *Sensors*, vol. 10, no. 6, pp. 5899–5926, 2010.

[112] T. H. Cormen, C. Stein, R. L. Rivest, and C. E. Leiserson, *Introduction to Algorithms*, 2nd ed. McGraw-Hill Higher Education, 2001.

Publications

International Journals

1. Vipin Pal, Girdhari Singh, R P Yadav, "Balanced Cluster Size Solution to Extend Lifetime of Wireless Sensor Networks," *IEEE Journal of Internet Things*, vol. 2(5), pp. 399-401, 2015.

2. Vipin Pal, Girdhari Singh, R P Yadav, "SCHS: Smart Cluster Head Selection Scheme for Clustering Algorithms in Wireless Sensor Networks," *Wireless Sensor Networks*, Scientific Research, vol. 4, pp. 273-280, 2012. ISSN: 1945-3086.

3. Vipin Pal, Girdhari Singh, R P Yadav, "Energy Efficient Clustering Scheme for Wireless Sensor Networks: A Survey," *Journal of Wireless Networking and Communication*, vol. 2(6), pp. 168-174, 2012. ISSN: 2167-7328.

4. Vipin Pal, Girdhari Singh, R P Yadav, "Analyzing the Effect of Variable Round Time for Clustering Approach in Wireless Sensor Networks," *Lecture Notes on Software Engineering*, vol.1, pp.31-34, 2013. ISSN: 2301-3559.

5. Vipin Pal, K R Yadav, Girdhari Singh, R P Yadav, "Effect of Frame Length on the Performance of Clustering in Wireless Sensor Networks," *International Journal of Computer Science and Management Systems*, vol. 3(1), pp. 41-46, 2011. ISSN: 0975-5349.

6. K R Yadav, Vipin Pal, Girdhari Singh, R P Yadav, "An Efficient Load Balancing Clustering Scheme for Data Centric Wireless Sensor Networks," *International Journal of Communication Network and Security*, vol. 1(3), 2012. ISSN: 2231-1882.

7. Vipin Pal, Girdhari Singh, R P Yadav, "Energy Efficient Clustering Algorithm to Exploit Heterogeneity of Sensor Nodes to Prolong Lifetime of Wireless Sensor Networks," *IET Wireless Sensor Systems*, (Communicated).

International Conferences

1. Vipin Pal, Girdhari Singh, R P Yadav, "Analytical and Simulation Modeling to Analyze Reliability State of Wireless Sensor Networks," *3rd World Conference World Conference on Information Systems and Technologies (World CIST15)*, Portugal, 13 April, 2015. Springer Series Advances in Intelligent Systems and Computing, Vol. 353, ISBN: 978-3-319-16486-1.

2. Vipin Pal, Girdhari Singh, R P Yadav, "Genetic Algorithm to Optimize Cluster Head Selection for Prolonging Lifetime of Wireless Sensor Networks," *3rd International Conference on recent Trends in Computing 2015 (ICRTC-2015)*, India, 12-13 March, 2015. Elsevier, ISBN: 978-93-81583-67-8.

3. Vipin Pal, Girdhari Singh, R P Yadav, "Effect of Heterogeneous Nodes Location on the Performance of Clustering Algorithms for Wireless Sensor Networks," *3rd International Conference on recent Trends in Computing 2015 (ICRTC-2015)*, India, 12-13 March, 2015. Elsevier, ISBN: 978-93-81583-67-8.

4. Vipin Pal, Girdhari Singh and R P Yadav, "Network Adaptive Round-Time Clustering Algorithm for Wireless Sensor Networks," *2rd International Conference on Advances in Computing, Communications and Informatics (ICACCI-2013)*, Mysore, India, pp.1299-1302, 2013. IEEEXplore ISBN: 978-1-4673-6217-7.

5. Vipin Pal, Girdhari Singh, R P Yadav, "Cluster Head Selection Scheme for Data Centric Wireless Sensor Networks," *3rd IEEE International Advance Computing Conference (IACC-2013)*, India, pp. 330-334, 2013. IEEEXplore ISBN: 978-1-4673-4529-3

6. Vipin Pal, Girdhari Singh, R P Yadav, "Optimizing Number of Cluster Head in Wireless Sensor Networks for Clustering Algorithms," *2rd International Conference on Soft Computing for Problem Solving (SocPros-2012)*, Jaipur, India, 28-30 Dec., 2012. Springer Series Advances in Intelligent Systems and Computing, Vol. 236. ISBN: 978-81-322-1602-5.

7. Vipin Pal, Girdhari Singh, R P Yadav, "Genetic Algorithm Based Method for Analyzing Reliability State of Wireless Sensor Networks," *International Conference on Frontiers in Intelligent Computing Theory and Applications (FICTA-2012)*, Bhubneshwar, India, pp. 583-591, 22-23, Dec., 2012. Springer-Verlag Berlin Heidelberg AISC, Vol. 199. ISBN: 978-3-642-35314-7.

8. Vipin Pal, Girdhari Singh, R P Yadav, "Energy Efficient Cluster Head Selection Scheme: A Change in Node Death Scenario of LEACH for Surveillance Wireless Sensor Networks," *International Conference on Parallel, Distributed and Grid Computing*, Shimla, India, pp. 865-869, 6-8, Dec., 2012. IEEEXplore ISBN: 978-1-4673-2925-5

9. K R Yadav, Vipin Pal, Girdhari Singh, R P Yadav, "Variable Frame Slot Clustering Scheme for Data Gathering Wireless Sensor Networks," *2nd International Conference on Current Trends in Technology (NUiCONE 2011)*, Ahmedabad, India, pp. 1-5, 8-10, Dec. 2011. IEEEXplore ISBN: 978-1-4577-2168-7

10. Vipin Pal, Girdhari Singh, R P Yadav, "A Restricted Tree Based Clustering Approach For Data Gathering Wireless Sensor Network," *International Conference on Advances, Modeling, Optimization, and Computation (AMOC-2011)*, IIT Roorkee, India, pp. 850-855, 5-7, Dec., 2011.

11. Vipin Pal, Girdhari Singh, R P Yadav, "Sleep-wake-up Scheduling Algorithms for Wireless Sensor Networks: A survey," *International Conference on Emerging Trends in Robotics and Communication Technologies (INTERACT 2010)*, Chennai, India, 3-5 Dec., 2010.

Brief CV

Vipin Pal received the B.Tech degree in Computer Science and Engineering from Chhotu Ram State College of Engineering (now, Deenbandhu Chhotu Ram University of Science and Technology), Murthal, Haryana in 2005 and M.Tech degree in Computer Science and Engineering from Guru Jambeshwar University of Science and Technology (GJUST), Hisar, Haryana in 2008. He is currently pursuing PhD degree in Electronics and Communication Engineering at Malaviya National Institute of Technology (MNIT) Jaipur. His areas of interest are Wireless Sensor Networks, Wireless Ad-Hoc Networks, and Algorithm Design. He is currently Graduate Student Member of IEEE, IEEE - Communication Society and International Association of Engineers (IAENG).

. . .