CLUSTERING AND CHANGE POINT DETECTION IN WIRELESS SENSOR NETWORKS FOR FIRE DETECTION APPLICATION

A THESIS SUBMITTED IN FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY
IN
COMPUTER SCIENCE & ENGINEERING

By

URVASHI CHUGH
Enrollment No: 116205

Under the Guidance of
Dr. Yashwant Singh

Department of Computer Science & Engineering and Information Technology
JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY WAKNAGHAT, SOLAN (H.P.), INDIA-173234

APRIL – 2016

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY
(Established by H.P. State Legislative vide Act No. 14 of 2002)
P.O. Waknaghat, Teh. Kandaghat, Distt. Solan - 173234 (H.P.) INDIA
Website: www.juit.ac.in
Phone No. (91) 01792-257999
Fax: +91-01792-245362

APRIL, 2016

CANDIDATE'S DECLARATION

I Urvashi Chugh certify that the work embodied in this Ph.D. thesis entitled, "Clustering and Change Point Detection in Wireless Sensor Networks for Fire Detection Application" is my own bonafide work carried out by me under the supervision of Dr. Yashwant Singh at Jaypee University of Information and Technology Waknaghat. The matter embodied in this Ph.D. thesis has not been submitted for the award of any other degree/diploma.

I declare that I have faithfully acknowledged, given credit to and referred to the research workers wherever their work has been cited in the text and the body of the thesis. I further certify that I have not lifted some other's work, paragraph, text, data, results, etc. reported in the journals, books, magazine reports, dissertations, thesis etc. or available at websites and included them in this Ph.D. thesis and cited them as my own work.

Urvashi

Urvashi Chugh

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY
(Established by H.P. State Legislative vide Act No. 14 of 2002)
P.O. Waknaghat, Teh. Kandaghat, Distt. Solan - 173234 (H.P.) INDIA
Website: www.juit.ac.in
Phone No. (91) 01792-257999
Fax: +91-01792-245362

APRIL, 2016

CERTIFICATE

This is to certify that the thesis entitled, "**Clustering and Change Point Detection in Wireless Sensor Networks for Fire Detection Application**" which is being submitted by **Urvashi Chugh** in fulfillment for the award of degree of Doctor of Philosophy in Computer Science & Engineering by the **Jaypee University of Information Technology**, is the record of candidate's own work carried out by her under my supervision. This work has not been submitted partially or wholly to any other University or Institute for the award of this or any other degree or diploma.

12/02/2017

Dr. Yashwant Singh
(Supervisor)
Associate Professor
Department of CSE & IT
Jaypee University of Information Technology,
Waknaghat, Solan, INDIA-173234

ACKNOWLEDGEMENT

Though only my name appears on the cover of this thesis, a great many people have contributed to its production. I owe my gratitude to all those people who have made this thesis possible.

Words are not sufficient but here trying to taking few words; I present my thanks to those people who were associated with me during this Ph.D. journey. Calling this journey as many faces; ups and downs come and goes but God was here with me and has given me support and power to come up at end of this work.

I express my sincere gratitude to Prof. S. C. Saxena, Vice-Chancellor, JUIT Waknaghat and Prof. Dr. T. S. Lamba, Dean Academic and Research, JUIT Waknaghat for their kind support extended to me for the successful completion of this research work. My deepest gratitude to Prof. S.P. Ghrera & all DPMC members who motivated me to improve better always.

I place on record unbounded gratitude to my teacher and supervisor, Dr. Yashwant Singh, Associate Professor, JUIT Waknaghat, for showing me the real path of sincerity and dedication. I have been amazingly fortunate to have an advisor who gave me the freedom to explore on my own and at the same time the guidance to recover when my steps faltered. He taught me how to question thoughts and express ideas. I am grateful and indebted to Dr.Yashwant Singh for allowing me to work under his supervision, and for his invaluable painstaking efforts taken towards my study.

I express my deepest sense of gratitude and sincere thanks to Prof. Ramana Murthy, Chairman Computer Science and Head Dept. of Mathematics, Osmania University, Hyderabad for his kind support. I want to present sincere thanks as he stood by me as God for me without any conditions. His patience and support helped me overcome many crisis situations and finish this thesis.

I am deeply indebted to my parents, friends and colleagues beyond all my family especially to my husband Amit Chugh for continuously motivating me with love. In lasts my loving thanks to my little champ Arnav Chugh who makes me learn how to remain busy with your work all time day and night does not matter.

Place: Waknaghat
Date: 11/5/16 **(Urvashi Chugh)**

TABLE OF CONTENTS

Contents	Page No.

Contents	Page No.

CHAPTER 3

MUTUAL EXCLUSIVE DISTRIBUTIVE CLUSTERING

CHAPTER 4

MUTUAL EXCLUSIVE HYBRID ENERGY EFFICIENT DISTRIBUTED CLUSTERING

Contents		Page No.

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 INTRODUCTION

1.1 Introduction

Wireless Sensor Networks (WSNs) are gaining wide-reaching consideration in recent years. In this chapter we have discussed the introductory concept of Wireless Sensor Networks. Section 1.2.1 explores history of Wireless Sensor Networks. Further design of wireless sensor network is presented in section 1.2.2. Section 1.2.3 discusses how wireless sensor networks are different from Adhoc networks. Section 1.2.4 and 1.2.5 presented applications of wireless sensor networks and few more concepts. Section 1.3 introduces concepts of clustering, its advantages and design issues etc., this section end with concepts of information fusion and change point detection. Section 1.4 is about genesis of problem which is formulated in section 1.5. Section 1.6 is about objectives of this thesis work. Approaches used are discussed in section 1.7. Contribution of thesis is pointed in section 1.8 and last section 1.9 will present thesis layout.

1.2 Wireless Sensor Network

Wireless Sensor Networks are constituted with tiny electro mechanical devices i.e. sensors nodes. Sensor nodes communicate via RF signals with one or more powerful sinks called base stations (BSs). Communication can be called single hope if they can contact directly to base station; or can be Multihop communication in which intermediate nodes takes part for communication from sensor node to base station. Formally sensor networks on basis of their functionality are classified into two types one is proactive and the other is reactive networks. Proactive networks are passive in nature and well suited for data aggregation type applications. In this type of networks nodes sense and send data on periodical basis i.e. on regular intervals. In Reactive networks, opposite to passive networks sensor nodes respond immediately and only to changes in the relevant parameters of interest. Reactive networks are more suitable for time critical applications. To prolong lifetime of sensor networks it will be more efficient if sensor nodes can collaborate with each other. That is the reason why sensor networks need management planes. As per requirement three management planes will work which are named as power management plane, mobility management plane, and task management plane [1]. These planes solve objectives like; first to make sensor nodes work collectively in a power efficient way, second to perform routing of data in wireless sensor network, and to share resources among them. Without these three planes, each sensor node will act and work as an individual. Power management plane ensures to manage how should a sensor node uses its power like when to turn off receiver; when to broadcast low power message and to quit from routing so that sensor nodes can save energy for sensing. The mobility management plane is in use when sensor nodes are not stationary in networks. It detects and registers the movement of sensor nodes. Mobility plane helps to maintain a route back to the user, and also sensor nodes could keep track of their neighbour sensors. The task management plane divides and schedules the tasks among sensors which are distributed in monitoring region. It is not always required that all sensor nodes keep sensing simultaneously in a specific region. It is also possible that some sensor nodes perform the task more than the others depending on their power level

1.2.1 History of Wireless Sensor Networks

History of WSN helps to understand evolution, requirements and tradeoffs in Wireless Sensor Network. Evolution of Wireless Sensor Network has occurred in military and industrial applications instead of civilian applications (environment surveillance, e-healthcare and agriculture) that are more common today. Sensor nodes have been developed during the Cold War. Sound Surveillance System is the first wireless network developed by United States of America on the bottom of ocean that provides likeness to the today’s Wireless Sensor Network during 1950s by the United States Military and used acoustic links for communication [2]. In late seventies, the United States Defense Advanced Research Projects Agency (DARPA) has initiated distributed sensor networks to investigate challenges during implementation of distributed sensor networks. These networks comprise of many nodes distributed spatially and operate independently but collaborate with each other to transfer information to the node that can use it in the best way. Distributed sensor networks have used Advanced Research Projects Agency Network (ARPANet) approach for communication. Components of distributed sensor networks have been identified during a workshop in 1978. Sensor networks include communication technologies, sensor hardware, software and processing algorithms. DARPA focused on military oriented research with SensIT program proposed during 1990s and designed sensor networks for battlefield monitoring and targeting [2]. Eventually, universities and governments start using Wireless Sensor Network in various real-life applications. After graduation, engineering students start using Wireless Sensor Network in the corporate world for industrial applications. These applications are based on heavy and expensive sensors. At that time, challenge has been to enable deployment of sensors in huge volume by minimizing deployment and maintenance cost as well as energy consumption per sensor so that sensor networks can be used in small civilian applications. Reduction in deployment cost whilst increasing functionality and scalability has been the major challenge [3]. Advancement in MEMS, digital electronics and wireless communication facilitate existence of modern Wireless Sensor Networks [99].

1.2.2 Design of wireless Sensor Networks

Each sensor node from WSN is packed unit of microcontroller, transceiver, power unit and battery [96] as shown in figure 1.1. Microcontrollers are like tiny computer. Sensor uses microcontroller who can have 8 to 128 pins. They usually run on 3.3 to 5 Volts and can works with frequency 8 to 20 MHz. They are having 256 bytes of RAM which is dependent on manufactures and on applications. Transceiver is used for communication between nodes and base station. Their two major functions are transmitting and receiving. Transmission media is usually RF signals but sometimes they also use license free range of communication 173, 433, 868, and 915 MHz and 2.4 GHz.

Sometime Transceivers have additional feature that they can posses any of these four states i.e. Transmit State, Receive state, Idle state, Sleep state. *Power* source of sensors is in form of batteries or capacitors. And last part is M*emory* unit for storage; usually sensors have storage in KB, Flash Memory, and RAM etc.

Sensor nodes are divided in four generations on the basis of obtrusiveness levels. These are named as Obtrusive, Parasitic, Symbiotic and Bio hybrid [4]. *Obtrusive* devices are large in size and weight like a shoe box. They are portable and capable to be a source of nuisance examples include sensor those are fitted on wearable camera for body tracking and halter electro-cardiographs. They are limited with high power dissipation.

Parasitic nodes have physical volume in few cubic centimeters, and weight is in tens of grams, but their physical parameters like size, weight and structure etc do not pretense them for normal behavior. Examples include sensors those are used in bio-metric watches.

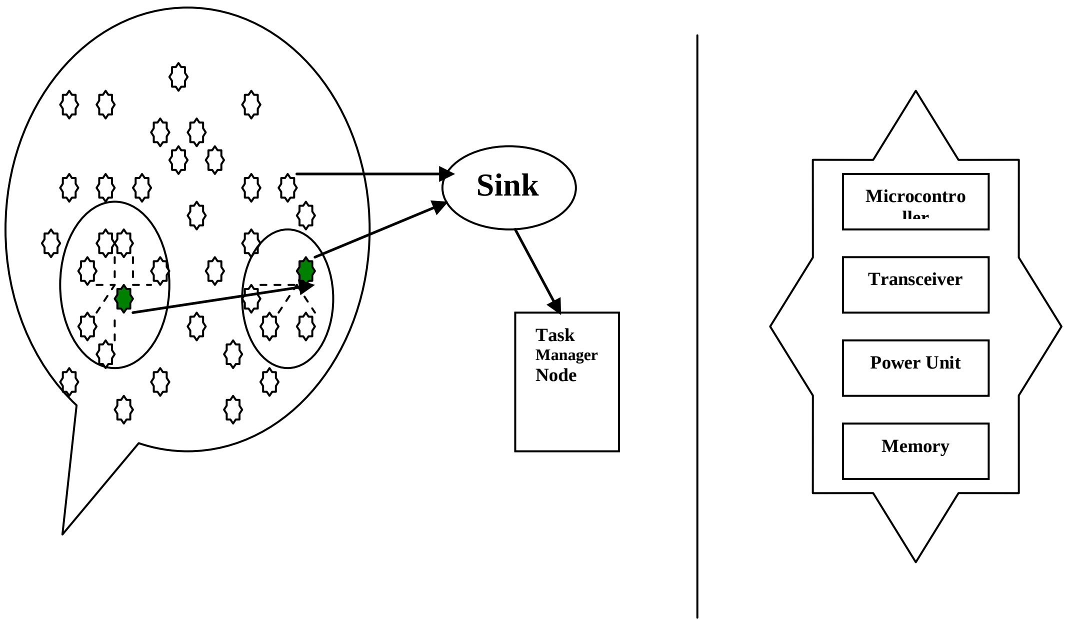


Figure 1.1- Wireless Sensor Networks

The power consumption of these nodes must not be larger than a few Millie watts. Next generation of sensor nodes is Symbiotic. These devices are called symbiotic as they posses well-matched benefit relationship with the target organism. These nodes are having size in cubic millimeter also termed as smart dust. Theses may used in numerous new in-body biomedical-monitoring applications. The technical faces of these devices are firstly their realization like autonomously-powered nodes; secondly they face challenge of safety requirements. The last i.e. forth generation is Bio-hybrid. The size of these devices is about a few cubic microns based on nanoscale molecular engineering. These devices can operate autonomously, as they get power due to chemical reactions generated by our biological systems.

1.2.3 Comparison of Ad hoc and Wireless Sensor Networks

Concept of Adhoc networks initiated during early 70s from United States military. Adhoc Networks and Wireless sensor networks have similarity that they both are decentralized in nature [89]. It a fact that ad hoc and sensor network have some more similar features like infrastructure-less, limited battery life, wireless channel for communication and self-configuration but they are having some major differences also [1,5]. Here we are differentiating these two networks on application basis, on resources basis that they may have

- Application: Major applicability difference between Adhoc network and in wireless sensor network is that Ad hoc network requires direct interaction with humans in contrast Sensor network involves applications that don't focus on human interaction but require lots of environment interaction such as Environment surveillance, E-healthcare, Railway monitoring, Nuclear reactors, Fire detection. Domain area of ad hoc network is different from the sensor network like Mine site operations, robot data gathering and urgent business meetings are some of the applications of ad hoc networks.

- Resources: Wireless Sensors networks are restricted with limited resources in contrast to ad hoc networks because of sensor's limited size and inexpensive nature. Sensor nodes possess limited computing power, restricted transmission power, low data rate due to small bandwidth and most importantly limited energy powered battery. Ad hoc network also has limited battery but it is an important factor in ad hoc networks not primary factor [87,88].

- Number of nodes: Numbers of nodes and density of network are order of magnitude higher in Wireless Sensor Network in comparison to ad hoc networks. Nodes are used in massive number because sensor networks are often deployed in hostile and unattended locations so must be fault tolerant networks. Sometimes, number of nodes may scale up to thousand.

- Failure rate: Sensors networks are more faulty and unpredictable because sensors nodes are fragile in nature. Moreover, sensor networks have very limited battery power and more susceptible to physical attack. Furthermore, surrounding environment may influence sensing component of nodes to transfer wrong readings. Sensor nodes have approximately 2 to 30 percent more packet losses than ad hoc networks.

- Nature of nodes: Laptop devices, Personal Digital Assistants, etc are used as nodes in ad hoc networks.

- Unattended locations: Sensors are deployed for longer periods of time with small batteries in remote locations without human intervention.

- Frequent topological changes: Sensors are more prone to topological changes not due to their mobile nature like in ad hoc networks but because of their hostile locations and limited battery power. Nodes die quickly either due to exhaustion of battery or physical attacks. New nodes are added in the network to replace old nodes or extend geographical region of the network.

- Mode of communication: In case of Wireless Sensor Networks, sensor nodes use broadcast mode for communication because sensor networks are very large in number. Ad hoc network possesses small number of nodes and follows point-to-point communication among nodes.

- In-network processing: In Wireless Sensor Network, nodes process the sensed data to reduce redundancy and extract more meaningful information before forwarding it further but in ad hoc network nodes communicate with original bits.

- Global identifier: Global identifier (ID) assignment is not recommended for Wireless Sensor Network because it induces huge overhead due to large size of network and rapid membership changes. Ad hoc network uses unique ID for its nodes.

- Data centric approach: In ad hoc networks, addresses are associated with every node because sensors follow node-centric approach. In sensor network, data sensed by the sensor nodes about a certain phenomenon is more important than address of sensors. For example, during water quality surveillance, user is interested in determining whether the value of turbidity is above a certain threshold. Therefore, instead of address sensors focus on data i.e. every action of sensor networks is based upon the gathered data.

- Cost: Sensor nodes are inexpensive due to small size and fewer resources in comparison to nodes deployed in ad hoc networks.

- Environment driven: Wireless senor networks are driven by the surrounding environment. Sensor networks generate new values whenever some changes occur in the environment. In sensor networks, pattern of traffic changes considerably on periodical basis. Their main objective is gathering of information from the surroundings whereas ad hoc network focuses on distributed computing.

- Data redundancy: In sensor networks, nodes are deployed with high density. So, nodes deployed in nearby locations measure approximately same values about the environment. Removal of redundant information helps in increasing energy-efficiency and network lifetime. Further, changes in environment occur very slowly so successive readings contain related values about the parameters.

- Self-organization: Sensor nodes must operate autonomously so that sensors can self-organize and self-heal under different circumstances.

- Communication range: Sensor nodes have small communication range only a few meters whereas range of ad hoc networks is in hundred of meters.

- Bandwidth: Wireless Sensor Networks communicate with broadcast so bandwidth is not a very major issue for them. However, nodes in ad hoc networks perform point-to-point communication over a shared communication channel so ad hoc networks are bandwidth deficient. Moreover, multiple transmissions over same frequency also result in interference.

Above mentioned differences clearly state that security proposals presented in the research papers by Candolin, Crepeau [6, 7], etc. for ad hoc networks cannot be used in Wireless Sensor Networks.

1.2.4 Applications of Wireless Sensor Networks

WSNs have applications in various fields like medical, military, weather forecasting and many more [69, 77, 81, 86]. In Table 1.1 we present application along with which type of sensor nodes used for them. The network size of wireless sensor network may change according to application from a small number of nodes to hundreds of sensor nodes or even more. Number of nodes contributing in a wireless sensor network is primarily depends on requirements associated to network connectivity, coverage and size of the region of interest.

Another important point of discussion in WSN is lifetime of sensor networks. The requisite lifespan of a sensor network may be some hours or several years, as per application requirements. The requirement is actually not only lifetime also degree of energy efficiency and robustness of the nodes is required. Other requirement includes a sensor network must support certain quality-of-service (QoS) aspects such as real-time constraints robustness, tamper-resistance, eavesdropping resistance, unobtrusiveness etc. When there are large number of sensors associated to network, it is practical impossible to recharge batteries of all individual sensor nodes and also it is very tough to ensure performance of wireless sensor networks. It will not be easy to support of time bound or mission critical data in challenge of limited energy. Now a days sensor network has also taken a turn towards ubiquitous sensing [85]

1.2.5 More about WSN

To enhance the lifetimes of the wireless sensors, design should be oriented to minimizing both computational and communication energy requirements and also all phase of the sensor systems should be efficient in energy perspective [78, 102]. Sensor network

Table 1.1- Applications of Wireless Sensor Networks

APPLICATIONS		
Application Areas	**Types of Sensor used**	**Examples**
Environmental [89]	Symbiotic	• Precision Agriculture [101] • Forest fire detection • Bio complexity mapping of the environment • Disaster management • Flood detection • Pollution detection • Habitat Monitoring [82
Military[98, 100]	Obtrusive	• Monitoring friendly forces, equipment and Ammunition. • Reconnaissance of opposing forces, terrain and Targeting • Border protection and security surveillance • Battle damage assessment • Nuclear, biological and chemical attack Detection and reconnaissance • Battlefield surveillance
Health	Bio-hybrid Symbiotic	• In-body bio-monitoring applications[97] • monitor physiological parameters of patients such as heartbeat or blood pressure • Telemonitoring of human physiological data • Tracking and monitoring human beings inside a hospital • Drug supervision done in hospitals
Home	Symbiotic Parasitic	• Home automation • Smart environment • Appliances • Entertainment
Industry	Parasitic Obtrusive	• Environmental control in office buildings • early detection of possible threats • Managing inventory control • Industrial and manufacturing automation • Distributed robotics
Other commercials [83]	Obtrusive	• Interactive museums • Air traffic control • Imaging sensors can identify, track, and measure the population of birds • Vehicle tracking and detection

protocols must focus first and foremost on power conservation. Also wireless sensor networks face a second challenge which is result of high node densities is that, the size of collected data will be large, and it will be very tough to store this data and to process it. Sensor nodes carry limited, generally irreplaceable, power sources. So it can be concluded that biggest challenge in WSN is limited battery power of sensors. With this limited battery power, sensor nodes will performs three tasks which are sensing, representing the sensed information and third is Transmitting/ Receiving. Out of these three transmissions is most energy consuming task.. Survey says that transmitting k bits to a distance d, the power requirement is proportional to d^2 or d^4 depending on free path or multipath. There are various theories, methods; inventions are in practice to reduce energy requirements so that lifetime of sensor nodes can be enhanced. Digital signal processor (DSP) is one of the proposed plans to make networks more energy efficient. DSPs are energy efficient processors and can do the analysis of sensor data locally. DSP demands some point like an energy-aware node should be able to adapt energy consumption when energy resources of the system diminish or as performance requirements change. Energy awaked nodes will results longer node lifetimes and more efficient sensor systems. Another characteristic in wireless sensor networks is required that all nodes should be homogeneous and also should have the same architecture. One study says cross-layer protocol design can support adaptivity and also can provide optimization across multiple layers from protocol stack. Cross layer design is very important to convene upcoming application requirements, especially when sensor nodes are limited by energy resource. Energy is always a sensitive area in WSN. Energy also affects sensing coverage along with lifetime of sensor networks. Some research is also focused on distribution of sensor nodes to optimize energy of sensor nodes. Distribution can also improve network lifetime. 2D Gaussian distribution seems beneficial over random distribution. One more energy saving technique is named as Clustering was introduced firstly by author Lin in 1997 [8]. The cluster based architecture was proposed and at that time it was stated that energy will be saved in this type of architecture.

1.3 Clustering

Clustering works for reduction in transmission cost and to save battery life in WSNs. Clustering is a grouping method in which, in place of a group of sensors only the group head (cluster head) will transfer the aggregated information. It's an implementation of hierarchical network. Hierarchical network in contrast of flat network having some more valued nodes or it can also be said at higher level nodes are having some extra duties. Same hierarchical approach is being used in clustering. Cluster Heads are at higher level and they collect sensed attributes from cluster members which are at lower level in contrast of cluster heads. Cluster heads will aggregate all sensory attributes and forward this aggregated information to the higher leveled cluster heads or base station (which may be sometime sink node).

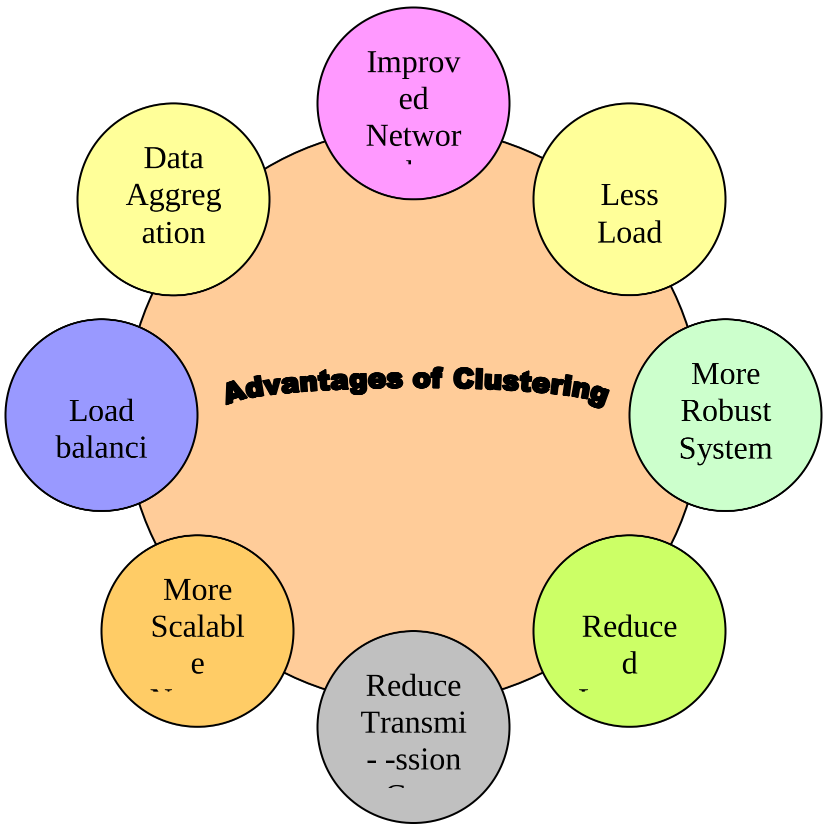


Figure 1.2- Advantages of Clustering

The primary objective of clustering is to minimize the power needed for communication to save battery life of sensors network [9] beyond which it's also having more advantages listed in figure 1.2. Clustering protocols can be compared on various features like centralized, distributive [59]. Some clustering protocols are power based, some may some under location aware protocols and some protocols have factors feature of multilevel clustering or some may support and multi-hop inter-cluster communication.

Centralized clustering algorithms have main role of base station it select cluster heads from sensor nodes. Cluster head will be decided by base station. Base station is having whole responsibility of timely cluster heads rotation as well. Centralized clustering protocols are benefited as message passing is not required in it so no extra overhead for network. Base station will broadcast message to inform which sensor nodes will act as cluster heads example of these typed protocols are LEACH-C and SHORT etc. The only disadvantage of centralized protocols is, need of high powered, efficient base stations. In *Distributed algorithms* sensor nodes elect among them which sensor suitable to be the cluster head. Cluster head is decided on the basis of message passing between nodes. Cluster heads are updated after next iteration of protocol. Examples of Distributed protocols are LEACH, EEUC, RRCH, ERA and DHAC etc.

Power base clustering algorithms consider residual battery life of sensor nodes or remaining network life time. Special roles and responsibilities for chasing cluster head will be decided on residual energy factor. Sensor nodes which are finding suitable with highest residual energy will be chased as cluster heads. Examples of some power based protocols are HEED and TEEN etc. *Multi-hope inters cluster communication* termed for path traversed for sending sensed information. Scenario in which sensor node is sending information directly to cluster head it is termed as single hop in contrast multi-hop inter-cluster communication is when sensed information will reach to cluster head via in between nodes. HEED and EEUC protocols also support Multi-hope communication.

Location awareness in wireless sensor networks is the ability of sensors with the help of GPS to know the location of all the other nodes. Sensor nodes are fitted with GPS device etc to find the exact location parameters. PEACH protocol is location aware protocol. *Multilevel clustering* shows the hierarchy in cluster heads means cluster heads are having their heads again. Multilevel clustering shows hierarchy in cluster heads. A cluster head is member of higher leveled cluster under some cluster head. Performance of clustering protocols are captured on following parameters

- *Time to network partition*- is defined as time at which the first node runs out of energy or can say dead, the network within a cluster is said to be partitioned. Dead node will reflect that some routes from network has become invalid
- *Average Lifetime of sensors*- is the second metric, which also guide about network life time as like time to network partition.
- *Average delay per packet*- is termed as the average time which a packet will take to reach at gateway (base station) started from a sensor node. Primarily energy is considered as vital factor in sensor networks, but few real time decisive applications demand that sensory data should be reported with least delay.
- *Network Throughput*- is termed as count of packets received at base station divided by the time taken for simulation. This is directly measurement of network traffic efficiency contributed by each cluster. Greater value of throughput directly means that network giving better routing for packets.
- *Average energy taken by packet*- This is also a metric to check performance of clustering protocols; reducing the energy in communication will results better network energy savings.
- *Average Power Consumed*- is defined as, average power consumption for message transmission. It is measured by taking values at different time span during simulation and averaging these values. It can be understood as the power consumed by message traffic in network.
- *Standard Deviation of Load per cluster*- Standard deviation of load is tested for various sensors deployments. Method to measure standard deviation of load is by changing no of gateways and also by growing sensor count in the network system.

There are several issues for clustering which are very important to consider in designing a clustering protocol. Issues need to cover for an efficient clustering protocol which could benefit network with enhanced lifetime, balanced load, scalable and with more robustness. The design issues that need to be considered are listed in figure 1.3.

Once sensors are organized into clusters they can contact to cluster heads by single hop or multi hop communication. In multi hop communication nodes which are near to a cluster head, will act as relay transfer the information to cluster heads and have higher load in contrast to other nodes. On the other hand in single hop communication the nodes which are at large distance from a cluster head have the maximum energy burden as they require more energy for long range communication need to reach at cluster heads. The cluster heads act as the first point for fusion in the network. Cluster heads perform first step data aggregation. They collect the packets from all nodes from their cluster, and aggregate the sensory data to produce a single packet. The size of packet after aggregation is always fixed irrespective of sensor node count in cluster i.e., it's not a matter how many number of packets were aggregated during data fusion.

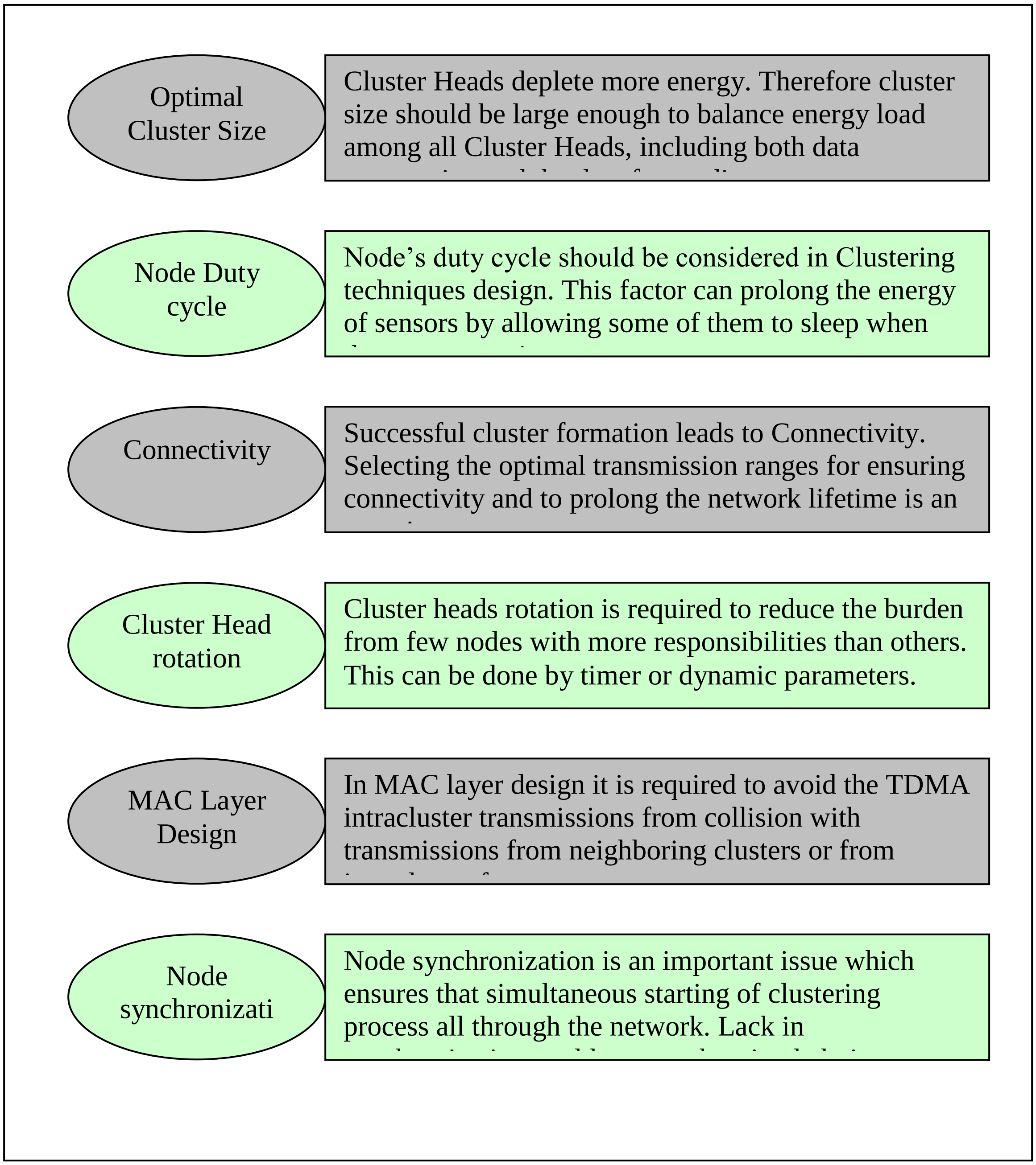


Figure 1.3- Issues in Clustering Protocols

Information Fusion is meant by collection of data and representing it in usable form. Information fusion results most accurate data from sensory values of any, at minimum cost. The idea of Information fusion is to achieve better information by combining different sensors output [10]. Information fusion system shows resolution, accuracy, speed, intelligence, insight, pragmatism etc. One more term in wireless sensor network is associated with information fusion is Data Aggregation. Information Fusion process is collection of firstly data aggregation and secondly filtering process. Data aggregation technique is used to combine several correlated data signals received from different nodes; which results effective data in form of smaller set [70]. Benefit of aggregation is reduction of data to be communicative to base station from cluster heads. It is concluded that simple aggregation techniques can also be used to reduce the overall data traffic to save energy. Information Fusion can be classified on three basis out of which first is *relationship among the sources* second is *level of abstraction* and the third classification is on the basis of *input and output*. According to the classification i.e. on relationship among the sources, information fusion is classified into three types named as complementary, redundant, or cooperative. Level of abstraction classifies information fusion into four levels which are signal, pixel, feature, and symbol. The third classification criteria i.e. input and output; divides information fusion into five categories which are as follows Data In–Data Out, Data In–Feature Out, Feature In–Feature Out, Feature In–Decision Out, Decision In–Decision Out [11].

Information fusion can solve multiple objectives like inference and classification also in, estimation, in feature maps, in abstract sensors, in compression etc. These objectives use different methodologies like fuzzy theory, neural network, abdicative reasoning, maximum likelihood, least square method and Kelman filter etc. A number of architectures and models have been proposed to supply guidelines for designing Information fusion systems. Three main categories of information fusion models are; Information-Based Models, Activity based models and Role based models. One very famous model name as Joint Directors of Laboratories (JDL) model from category Information based models. JDL's terminology has applicability in defense; it shows the level 0 to level 4 refinements of inputs taken from sensor nodes. Level 0 starts from source preprocessing moves to level 1 which is purposed for object refinement and then after level 2 perform situation refinement after which output of level 2 proceed towards level 3 for Threat refinement and then last level of refinements is level 4 which is for process refinement. Many more information fusion models are working like one is Dasarathy Model. Drasarthy model is also known as Data-Feature-Decision (DFD) which works in grained approach, it fuse elements on their inputs and outputs basis. The two activity-based models named as OODA (Observe, Orient, Decide, Act) and second is Intelligence Cycle. These models have applicability in various domains. It is found in study that in WSNs data on communication is directly dependant on information fusion. More data to communicate means more power required. The challenging part is limited energy of current sensor nodes, due to which it is usually required to reduce the overall data traffic in networks. Change point detection further reduces data communication requirements in contrast of regular data gathering. Regular Data gathering keeps transmission active all time but in change point detection data will be communicated only at time of particular event detection.

Change detection in WSN plays a very important roll in multiple fields like in intrusion detection, fault diagnosis and also in monitoring applications [12]. Change point detection is finding those dimensions which don't follow normal pattern as like general data. Change point detection is also termed as monitoring of sudden changes in the regular parameters of data. These abrupt changes in WSN can be observed by either sensed data results or traffic-relevant dimensions in the network. Such abnormal behavior can occur due to many of reasons out of which one can be malicious attacks; also it is possible that intrusions on a network can produce change in sensory measurements. Other reasons like faulty sensors or unusual phenomena could also result changed outputs. Change point detection is having its applicability in security fields too; like in Intrusion detection. Change point detection can be used in sensor networks to detect unusual changes in the monitored environment. Networks can also raise an alarm when such type of abrupt changes or misbehaviors occurs in networks. Some authors use term Outlier in place of change point detection. The Author Yang Zhang has presented classification, identity, degree and different ways of outlier (change point detection) [13]. Four parameters can classify outlier among them, first is *Input Sensor Data*, second is *Type of Outlier* third is *Identity of Outliers* and last is *Degree of Being an Outlier*. The first parameter Type of input sensor data resolve that which outlier detection method can be opted for analyzing this data i.e. data on attribute basis or correlation basis. Second parameter, types of outlier divide outlier into two types may be either local or global, depending on the scope of data used for outlier detection. Third factor is identification of outlier; there are mainly three causes of outliers occurred in network one is outlier can be due to noise and errors, second is outlier can be due to an event, and one more cause of outlier can be malicious attacks. Outlier detection is not only limited to find out which data is not following normal pattern like generous sensor data, but also it provides certain techniques to calculate the degree of deviation of data measurements from normal pattern. Outliers are measured in two scales, i.e. scalar and outlier score. Several techniques are researched for detecting change in sensed inputs. Some of these are based in *Statistical-Based Approaches;* some are based on *Nearest Neighbor-Based Approaches* etc. The author Yang Zhang has provided a relative table which is very useful to select a method appropriate for applications based on these characteristics like data type, outlier identity, and outlier type and outlier degree.

1.4 Genesis of Problem

Wireless Sensor Networks (WSNs) are gaining widespread importance day by day due to its broad categories of applications in every field; from home applications to defense surveillance monitoring; from personal vehicular applications to medical field applications etc. As most applications require a long network lifetime but challenges like tight energy constraints, high density WSNs are critical. In case of high density WSN will turn off few nodes but remain operational as well. The point to note is that WSNs have touched every field in applicability therefore it needs to fulfill multiple design objectives as well like Small Node Size, Low Node Cost, Self Configurability, Scalability, Reliability, Fault Tolerance, Security, Channel Utilization, QoS, Low Power Consumption, and Application Specific needs etc. Therefore, a WSN should be challenged to efficiently handle some of the trivial issues like power consumption, network life time and to prove themselves application specific like change point detection. Out of several issues limited battery power or to save battery power aimed to make network more sustainable is the crucial one. Many energy saving techniques were in discussion among them Clustering is most fancy one. Clustering techniques solve multiple objectives along with power economy. The centralized clustering protocols face disadvantage of high powered base station requirement. It is also found that distributive protocols work on residual energy factor but residual factor is not directly considered infect most famous protocol has also used this factor for probability calculations. Why not Distributive protocol can directly consider residual energy factor only? Which method should sensors follow to select cluster heads if residual energy is only criteria?

Clustering can reduce power consumption by reducing data on communication. It is also a well understood concept that event detection in WSN also has less data communication requirements. So merging these two concepts i.e. clustering and change point detection together can produce more sustain networks which can have better life time. Up to best of our knowledge no clustering protocol has done work on this. Work can be done on energy consumption and change point detection together.

1.5 Problem Statement

Wireless Sensor networks are challenge with many issues out of which we have targeted on two which are to save networks energy and to make network application specific. Sensor nodes use their battery for detecting and collecting information from the region of deployment but the hard fact is that there is very little scope for recharge or change batteries of sensor nodes. These sensing nodes collect and pass the information from the network towards the base station for further actions. We need to consider energy consumption as a major factor of concern for a better functioning and a longer life timed network. So it becomes very important to save energy of sensor nodes and also to use remaining energy in optimized way. If this thing would not be considered in that case network will face portioning problem and sooner become unresponsive.

We have hit on energy constraints and tried to find out solution by adopting which can save energy. For energy saving work is done on clustering. The motto of this research work is to propose clustering protocols that are energy efficient and those can support change point detection or event detection. In literature, many authors have contributed toward grouping of sensor nodes and routing the information in WSN for energy conservation purpose. But best of our knowledge none off clustering protocol has targeted on change point detection application as well. To tune network in such a way that it need to communicate only at event detection can result more energy saving. Event detection is also can be understood as “finding awry reading that deviate from normal pattern of observations.

1.6 Objectives

The following objectives corresponding to the problem statement have been achieved in this work.

a) To group the sensor nodes in WSN for reducing the power consumption in communication for improving network life span of WSN.

b) To design a new clustering protocol which should be energy aware by which energy efficiency could be improved.

c) To design and develop change point detection scheme for fire detection application in WSN.

Mapping these objectives onto WSN, the following three distinct intended deliverables have been achieved:

I) Mutual Exclusive Distributive clustering (MEDC) protocol proposed. This result following points

i) Cluster Formation

ii) Network lifetime improvement

II) Mutual Exclusive Hybrid energy Efficient Clustering (MEHEED) protocol proposed. This result following points

i) Cluster Formation.

ii) Include benefits of HEED protocol as well.

iii) Further Network life time improvement in contrast of MEDC.

III) Change Point Detection along with MEDC protocol (using Fuzzy Logic) which results

i) Fire Detection

1.7 Approach followed

To develop new clustering protocol distributive approach is firstly considered. Distributive approach leads to each sensor node to participate in cluster heads selection. Each sensor node will advertise residual energy. On based of message passing system nodes with highest residual energy will be selected as cluster heads in their range of communication. Based on this approach we proposed MEDC protocol.

Further to improve efficiency of MEDC clustering protocol HEED’s protocol logic was combined with it and a new protocol was proposed MEHEED. MEHEED takes advantage of MEDC and HEED. MEHEED will firstly select those sensors as cluster heads that are having probability equal to one without considering residual energy and without floating advertisements. In case if no sensor node under range of communication is having probability equal to one in that case MEDC logic will be used and sensors will decide their cluster heads to that sensor that is having highest residual energy. Advertisements will be float in this case only Benefit of MEHEED is it showed better performance in improvement of network lifetime contrast to MEDC.

For Change point detection we had chose fuzzy logic approach. Reason for choosing fuzzy logic is it can fuzzify crisp value to fuzzy sets of real scenario. After fuzzify input values fuzzy rules can map to inputs toward outputs. The point of using FIS (Fuzzy Inference system) is that it can show fuzzy output to crisp value after Defuzzification.

1.8 Contributions

(i) Mutual Exclusive Distributive Clustering (MEDC) protocol has been developed. The work assumes that each sensor node is having equal capabilities and they are having non rechargeable battery. MEDC protocol floats advertisements and decides cluster heads in distributive manner (each sensor participates in selection process). This protocol selects that node as a cluster head which have highest residual energy under that range of communication. MEDC protocol has shown significant improvement in network life time.

(ii) Mutual Exclusive Hybrid Energy Efficient Distributive Clustering (MEHEED) has been developed. MEHEED is merged solution of MEDC with most famous Hybrid Energy Efficient Distributive Clustering (HEED) protocol. MEHEED clustering takes advantage of both protocols and shown significant improvement. The reason for betterments of MEHEED is that it will try to select cluster heads firstly on probability basis. Message floating will be done only when there is no sensor node from range of communication is having probability equal to one

(iii) A fuzzy logic based change point detection technique for fire application has been developed, which detects the change at cluster head level. This approach can also have positive impact in saving energy; because cluster heads will transmit only when they found more then 50% chances of any event from aggregated inputs.

1.9 Thesis Layout

Thesis work is organized into five chapters. First Chapter is about introduction to WSN, Clustering and Change Point Detection. Introduction includes basic terminology related to these topics as well. Second chapter presents intense survey related to thesis work. Literature survey covers research history and trends of evolution of WSN. This chapter includes existing clustering protocols. We had tried to present tabular summery of clustering protocols on feature basis. Next part of this chapter presents literature survey under different change point detection methods.

Third Chapter is our first contributory work named MEDC protocol. This chapter presents introduction to MEDC protocol, its algorithm and its results under various parameters. We had presented examples to show working of MEDC protocol. Simulation for MEDC is done on MATLAB. For evaluation of this clustering protocol network lifetime criteria is taken, Results of MEDC are compared with HEED; and it is shown that MEDC is performing better than HEED. MEDC protocol has shown improvement in network life time. We had also shown performance of MEDC under various parameters. Forth chapter is about our second contributory work named MEHEED. MEHEED protocol is merged solution of two clustering protocol one is MEDC and the other is HEED. This chapter will first present concept of MEHEED clustering protocol then algorithm after that experimental results under various parameter. Performance of MEHEED is compared with MEDC and HEED as well.

Fifth chapter introduces Fuzzy Inference System (FIS). Then it will show our third contributory work i.e. change point detection for fire detection scenario. FIS will be input with three parameters and output will be fire probability. Thesis ends with chapter six which is about conclusion of thesis work and future work possible.

CHAPTER 2 RELATED WORK

2.1 Introduction

This chapter presents a comprehensive literature survey of Wireless Sensor Network, Clustering and Change Point Detection techniques. Section 2.2 includes research history of WSN along with timeline and trends of evolution in WSN. Section introduced how wireless sensor network evolved and what are design objectives. Section continues with survey with work of many researchers by which we came to know about layers in WSN and management planes. Study also enlightened that WSNs are different from Adhoc networks and how sensors comes after three generation. One more interesting point come up in survey that on researcher also proposed that Bluetooth can be communication method in place of RF signals. Section also includes with application of wireless sensor networks. Out of survey we also come to know about issues involved in design of WSN. Section 2.2.1 introduced different categories of wireless senor networks section 2.2.2 discussed different energy saving techniques. Many authors has discussed about issues out of these issue power management is crucial one. One author proposed DSP for energy efficiency and some other work proposed that Gaussian Distributions can also further improve power management to increase life time of sensor networks. A number of power management or can say energy saving techniques are reviewed out of which we found clustering is most fancy and advance technique for energy economy. So we take a next turn in survey for Clustering. Section 2.3 presents survey of clustering protocols. Many clustering protocols are proposed by authors. Authors have worked to find clustering objectives, benefits and issues. Later we surveyed first clustering protocol LEACH. LEACH was presented by author Heinzelman to solve energy efficient routing and data aggregation in smarter way. Later one more scheme PEGASIS tried to improve work done by LEACH but PEGASIS was a more type of Routing protocol. Survey continues TEEN protocol which was developed for reactive networks. Next part of survey include HEED protocol; and here a new design factor come up in light which was residual energy of sensor nodes along with probability. Next we come up an unequal cluster scheme EEUC. Further one more scheme PEACH was developed for both type of sensors location aware and location unaware. Survey includes more clustering protocols like SHORT, DHAC, and SEACH. In last of this section we have tried to summarize different clustering protocols on different parameters in tabular form. In this part we found no clustering protocol is updating clustering parameter as latest residual energy calculated. Moreover no clustering protocol talks about energy saving in event detection. Only TEEN protocol has considered about reactive networks but the authors has not defined by which scheme further nodes should report to change detection.

Next part of this chapter is about Change detection which was next point of attraction from where we had taken our third objective. Change detection is applied to various fields like in Habitat monitoring, Industrial monitoring and environmental monitoring. Health and medical field also have importance of change point detection. In WSN multiple change point detection schemes can work. We had reviewed change point detection by Fuzzy logic, Dempster Shafter Theory, Bayesian, Neural Networks and some other methods like Kelman Filter or nearest neighbor methods. These methods were previously known from conceptual point of view; on the other side by brilliant authors these methods got importance in the field of change point detection in WSNs. Section 2.4 has discussed gaps found in literature survey.

2.2 Wireless Sensor Networks

In year 2000 it was firstly published that Pico Radio can support wireless sensor networks for communication [14]. At that time it was just an assumption that in multiplicity of scenarios, these sensor and actuator networks might excel. As time proceeds in 2001 Intel Research Lab at Berkeley focused on WSN. Further in year 2002 two design projects came to study names as Spring 2000 Design Project and Spring 2001 [15]. These projects, has developed a design course by the Tennessee Technological University and the University of South Florida. These projects was aimed for making students study, design, fabricate, and test a wireless sensor system that transmit data over a radio-frequency link for centralized processing. In same year 2002 Akyildiz defined sensor networks as convergence of micro electro-mechanical systems technology. Paper had presented five layered architecture along with three management planes of wireless sensor networks [1] then after the algorithms and protocols available for each layer in the literature are explored Author has discussed wide range of applications in military, environmental, health, home and in other commercial areas. Paper discussed various factors which influence the design of sensor networks factors like scalability, mode of transmission, fault tolerance, hardware limitations, topology and power consumption. Survey also says that, sensor networks are different from Adhoc network so Adhoc routing protocols are not useful for WSN. The reason for this mismatches or can say applicability difference, is *energy constraint* of sensor networks. This is worth to understand that sensor nodes are limited in battery life. So it's is very important to concentrate on energy saving of sensor networks. Number of concepts for energy saving has been discussions; clustering is one of them. Author has discussed open research issues for the realization of sensor networks. In same year i.e. 2002 author Goldsmith [16] has presented wireless sensor networks as application of wireless Adhoc networks. Author has presented new and exciting applications of Ad hoc wireless networks, also significant technical challenges. Paper is overview of ad hoc wireless networks and their applications but emphasis is on energy constraints. Energy is considered as major issue for nodes. After that authors discussed advances in factors like the link, multiple access, network, and application protocols for these networks. Paper also presented cross-layer design of these protocols which should be suitable when energy is a limited resource.

Next in [17] author Chong, et al have presented history of Wireless sensor networks and research carried out during past years. Author has presented three generation of sensor nodes in terms of feature like weight, size and cost etc. Paper give review how sensor network started from Military Sensor Networks in 1890 and proceeded at next level in form of Distributed Sensor Networks to current 21st century. Paper has presented several new applications like infrastructure security, industrial sensing, environment surveillance and traffic control etc and also discussed technological trends and challenges that may effect on development of Wireless Sensor Network. Author reviewed that sensors networks faces challenges in data handing, data communication, and managing sensors; DSN also identified few of them. Sometime Because of environmental issues wireless ad hoc networks pose additional technical challenges like network discovery, network control and routing, collaborative information processing, querying, and tasking. The author concluded his work by providing some results in algorithms applied to sensor networks, which includes some localized algorithms.

Up to this time research was going for wireless sensor networks and it was taken that mode of communication is RF signals but in 2004 one more study come into light [18]. Author has implemented wireless sensor networks for security on a new communication method i.e. on Bluetooth. Bluetooth is a wireless technology used for short-range communication. This technology is benefitted with low cost, consume less power, which become reason why number of industries showing interest towards Bluetooth technology. Due to these benefits authors has accepted this technology for a WSN targeted for security. This paper includes the realization matter related to power management, what should be system structure, how to make networks self configurable so that they can handle errors and perform better routing. But this study do not get further valued and standards moves back to RF signals due to long range of connectivity. In 2004 paper [19] has introduced a new design for sensor network under different dimensions. Authors has conducted a workshop and studied that during past years, wireless sensor networks have found their importance in wide variety of applications for which they have discussed typical requirements regarding hardware issues and software support. In spite of problems in a multidisciplinary research area such as wireless sensor networks, where close collaboration is needed between users, application domain experts, hardware designers, and software developers for implementing efficient systems still they have justified their view by demonstrating specific existing applications which occupy different points in the design space. Authors highlighted the fact that the sensor nodes are independent to each other, while they have constraints of limited size, cost and energy. So it is right to say that varying size and cost constraints can directly lead to energy availability, computation power, also can have better storage also other resources. Authors have partitioned sensor nodes roughly into four classes based on their physical size, which are named as brick, matchbox, grain, and dust. Paper also discussed requirements of wireless sensor network shown in figure 2.1.

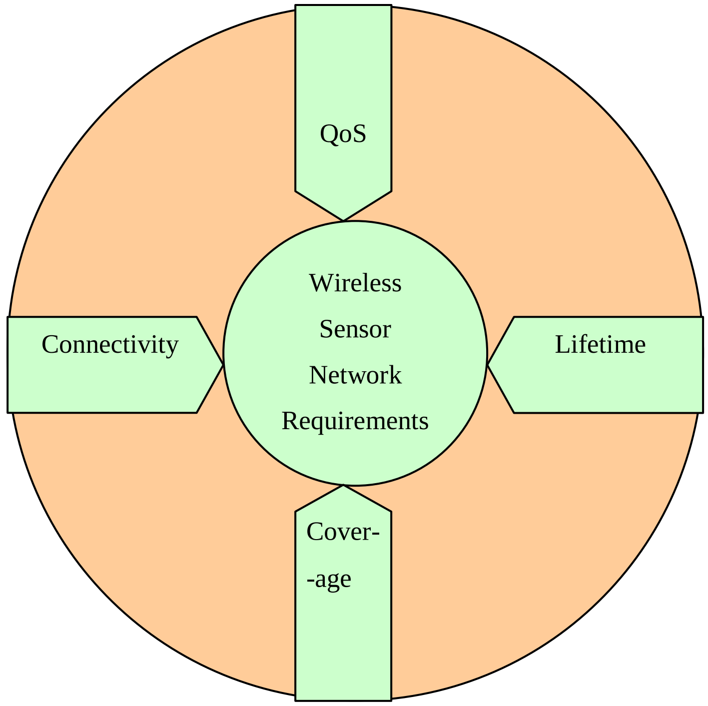


Figure 2.1- WSN requirements

In 2006 further sensor networks are categorized in four different generations listed first is obtrusive sensor networks, second category is parasitic sensor networks, third category of sensor networks is symbiotic and lastly sensor networks can also be classified as bio-inspired [4]. Authors have also discussed how WSN has evaluated under different trends. Their research has based on characteristics like what resources they are using, what are their energy level, how they are formed and their respective costs. Survey includes step by step generation of sensors how they start from shoe box size and how they reached at such a small size like a dust particle. Additionally, they discussed key research challenges in Wireless Sensor Network together with, node architecture, and hardware technology for future directions. Survey included how common themes have emerged. Authors have focused on technologies to reduce power consumption. In 2007 author Baronti had presented [20] a new standard IEEE802.15.4 name given by Zig Bee. IEEE802.15.4 is standard name given to Low-Rate Wireless Personal Area Networks (LR-WPAN).Authors had overviewed sensors hardware first then introduced 802.15.4 after which authors worked to present MAC layer and Network layer protocols for this standard along-with various performance characteristics for IEEE 802.15.4 and have concluded with the future trends of Wireless Sensor Network. Yick, have presented more recent literature survey of Wireless Sensor Network including key issues in Wireless Sensor Network at node level as well as at group level [21]. They have also surveyed recent issues like localization, network security, data aggregation, compression, coverage and presented a number of research issues for Wireless Sensor Network. Under networking one interesting topic data aggregation has used cluster approach of communication. Author karl has discussed protocol architecture of wireless sensor network in book [84]. Clustering is approach to save communication energy. We will discuss this topic in detail later. Further in 2009 author Burratti had done one more survey on WSN and presented different design requirements according to applications in [22]. Paper has presented a case study on environment monitoring. Standard IEEE 802.15.4 which was introduced in [20] motes is used for case study. Author has clearly shown that regular monitoring without panel uses more battery power; so battery drained out very fast. Here we also concluded that regular monitory uses more battery power then it can be said that event detection will definitely reduces power consumption. From here we get a new objective will be event detection instead of regular data gathering. In [23] authors have discussed challenges, design principles, and technical approaches of industrial wireless sensor networks (IWSNs). Authors have discussed about radio technologies, some techniques for energy harvesting, and also presented cross-layer design. The IWSNs have several advantages over traditional wired industrial monitoring and control systems like self-organization, rapid deployment, flexibility, and inherent intelligent-processing capability. They have created such industrial systems that posses maximum reliable and also posses self configuration property. These systems can respond to real-time events very quickly and can take proper actions. Work is done firstly to represent technical challenges and design principles of IWSNs. Author surveyed technical challenges like resource constraint, Dynamic topologies and harsh environmental conditions, quality of service requirements, data redundancy and Packet errors and variable-link capacity etc. The study also presents issues in hardware development, system architectures and protocols, and software development for IWSNs. Author also enlightens different design goals for IWSNs. In their study the paper also overviewed on some open research issues like what is the optimal deployment technique for sensors, what security techniques should be used, and how t make different IWSN interoperable?

2.2.1 Classifications of Wireless Sensor Network

Wireless Sensor Network can be classified on different categories. Here we want to present survey of different types of sensor networks refer figure 2.2.

Homogeneous or Heterogeneous

Homogeneous networks have all nodes of same type [24]. All nodes have same battery time same memory capacity and same computation power etc. Nodes will equal capabilities and resources are deployed over the region and duties are divided in between them without any priorities. On the other hand heterogeneous network have different type of nodes on region of deployment. Some nodes have extra capabilities like have more computation power more memory or more energy. These sensor nodes are deployed in same region with low prior nodes. These special nodes are deployed in planned matter to hand over some extra responsibilities over them [62].

Proactive or Reactive

Sensor nodes from proactive sensor networks keep sensing on regular matter. They remain active all the time and send data to base station on intervals. Reactive networks in contrast of proactive networks send information only when some uncommon event happens. Sensor nodes keep observing but transmit only when they detect any change in regular phenomena. Proactive sensor nodes have shorter lifetime as comparison of Reactive networks.

Structured or Unstructured

On infrastructure basis WSNs can be structured or unstructured. Usually in unstructured Wireless Sensor Network, a large numbers of sensor nodes are deployed without any strategy. We can also say in ad hoc manner when we are not very sure about the position of nodes in the geographic field. These sensor nodes operate in an unattended environment. These networks are tough to maintain and to satisfy QoS because there is no assurance of connectivity and failures detection is near to impossible. On the other hand sensors in Structured WSNs are deployed in a pre-defined manner. Deployments are planned in engineered way to satisfy and meet network requirements.

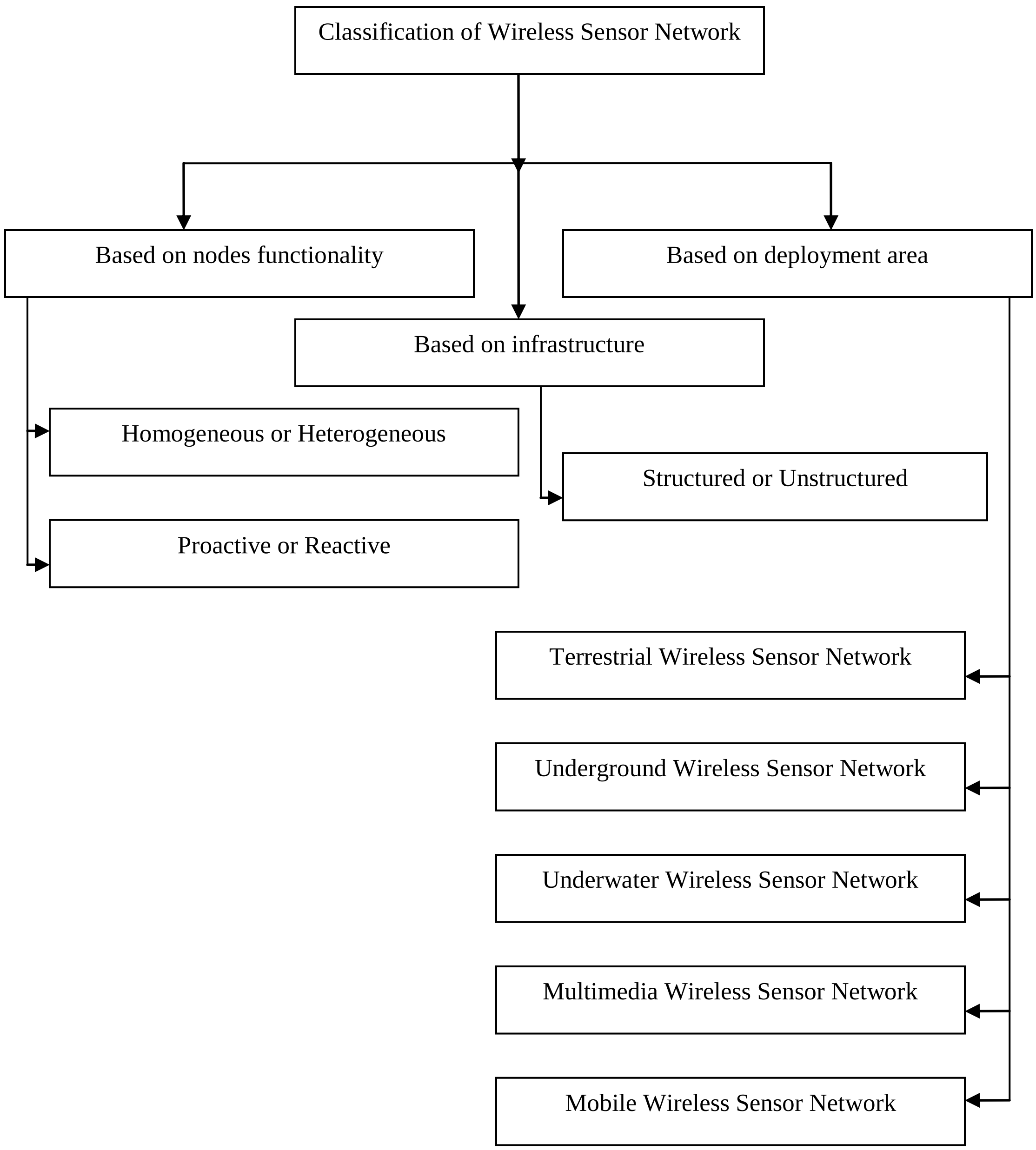


Figure 2.2- Classification of Wireless Sensor Network

There maintain is easy and cheap in comparison of unstructured WSNs. Structured wireless sensor network offers optimal coverage even with small number of nodes whereas unstructured Wireless Sensor Networks may still have some uncovered regions [21].

Terrestrial Wireless Sensor Network

Terrestrial Wireless Sensor Networks are generally deployed in a fixed region for monitoring purpose generally made up of hundreds to thousands number of nodes [1]. Monitoring can be done in surveillance field where army needed for security or can be in hospitals, in forests etc. Infrastructure of Terrestrial can be structured or unstructured. Usually sensors are deployed with fixed battery power which is most of time non rechargeable; some researchers also purposed solar cell to charge batteries of sensor nodes. To save energy conservation techniques can also be used in terrestrial networks. Sensor nodes can maintain low duty cycle operation, can perform aggregation can use clustering etc.

Underground Wireless Sensor Network

Underground Wireless Sensor Networks are termed when sensor nodes are deployed in caves, mines or below ground [25, 26]. These networks are aimed to observe underground conditions and their base station or relay nodes can also be on ground. Generally these networks are properly structured their deployment is planned. Underground networks have high cost of deployment; also equipment and maintenance cost of these are high in contrast of terrestrial networks. These networks are challenged to show quality of service in communication because of more attenuation in signals and to be reliable in transmission through soil, water and other present minerals. Battery life is also a challenging concept because it's totally impossible to charge batteries or to replace them.

Underwater Wireless Sensor Network

In Underwater Wireless Sensor Network sensors are deployed in water environment. Sometimes may be under deep water oceans sometime along with autonomous underwater transport [27, 28]. Deployment of networks in underwater environment costs expensive h. Mode of communication in underwater wireless sensor Networks are acoustic waves. Problem incurred with acoustic waves is limited bandwidth, high propagation delay and large attenuation. One major problem is also failure ratio is higher in underwater sensor networks. So underwater nodes must possess self-configuration property and adapt themselves to harsh conditions of underwater environment.

Multimedia Wireless Sensor Network

Multimedia Wireless Sensor Networks [29] is constituted with nodes that capture information from environment as multimedia contents like video, audio and images. Nodes in these types of networks are outfitted with microphones and cameras. Large number of nodes is restricted due to Expensive nature of nodes. Multimedia Wireless Sensor Networks demand huge bandwidth, high energy, high quality contents, processing and compression of large amount of data [29]. As multimedia contents like videos require huge bandwidth, So Wireless Sensor Networks require huge bandwidth, high quality contents, high energy, processing and compression of large amount of data [29]. Although it is a challenging task in multimedia Wireless Sensor Networks to guarantee high quality of contents because of varying amount of delay and channel capacity but certain level of quality must be ensured for reliable transmission of data. To minimize the redundant information and merging of similar information; Aggregation, compression and filtering should be used as a competent tool. This results in enhancing performance of the network.

Mobile Wireless Sensor Network

Mobile Wireless Sensor Networks is constituted by nodes which have the ability to move and gather data from the environment [21]. Mobile sensor nodes have the entire competence of static nodes like sense, compute and communicate. A basic distinction between static and mobile nodes is the capability of mobile nodes to relocate and rearrange themselves in the network. Mobile Wireless Sensor Network commence with initial deployment of nodes. Later on, nodes relocate themselves over the network to gather information. Mobile nodes also exchange information with each other whenever nodes come in communication range. Distribution of data is another feature that differentiates mobile sensor nodes from the other nodes. Static networks use fixed routing whereas dynamic routing is used in case of mobile networks. Some of the key challenges faced by mobile sensor networks are like Navigation, deployment, maintenance, coverage, energy dissipation and self-organization. Environment surveillance, target tracking and real-time monitoring of affected areas in case of disasters are some of key applications of Mobile Wireless Sensor Network. Manual deployment of sensors is not

2.2.2 Schemes for Energy Saving

DSP (Digital Signal Processor)

Energy constraint was very important issue for research in wireless sensor networks; in 2002 author Wang has proposed energy efficient digital signal processor for wireless sensor networks [30]. Paper is mainly focused on energy efficient wireless sensor networks design. Authors believe that in a wireless communication network of sensors a low-power DSP can perform local analysis on sensor data. Authors has also discussed that micro sensors network enable a variety of new applications like warehouse inventory tracking, location sensing, machine-mounted sensing, patient monitoring, and building climate control etc. This paper has considered one application that use micro sensor, and chose communication by acoustic waves for environmental monitoring. These types of sensors are adaptable in nature and have applicability in speech recognition applications also in medical diagnosis and in traffic monitoring. Further paper has presented Wireless Sensor Node Architecture. Under Architecture authors has discussed concept of clustering under Multihop communication. Clustering reduces power dissipation by grouping sensors and making cluster heads in groups. Authors had also worked on reduction of storage requirement by efficient networking protocols, and local signal processing.

2D Gaussian distribution

Paper [31] has overcome the issue of sensing coverage and also focused on battery power constraint. Authors surveyed that WSN are having battery power constraints which can affect Coverage and also have great impact on lifetime. The authors have presented framework using 2D Gaussian distributions and showed the improved coverage and also enhanced lifetime of a WSN. Paper has evaluated Gaussian dispersion on different dimensions to study coverage and lifetime of network. Author admitted different Gaussian parameters but follow uniform distribution. The study and identification of intrinsic properties of coverage/lifetime in terms of Gaussian distribution parameters is studied and results are obtained. Further authors determine different sensor deployment strategies, which could satisfy a well enhancement in lifetime. On the basis of analytical models author proposed two algorithms for deployment to increase network lifetime. These deployment strategies have complexity in polynomial time and give optimal solution. The work has clearly showed that Gaussian distribution can enhance lifetime of sensor networks. Network life time is very critical and important issue in each type of wireless sensor networks.

Distributed algorithm using Lagrangian

To enhance network lifetime one more paper [32] has presented one distributed algorithm in wireless visual sensor networks. To improve network lifetime author has used a strategy in which the source rates are optimized, does routing and use best encoding powers. Distributed algorithm was developed which was using Lagrangian duality to improve the lifetime of wireless visual sensor networks. Work is done after investigating both large-delay applications and small-delay applications. The proposed algorithm achieves better network lifetime in contrast to the network lifetime maximization algorithm for conventional wireless sensor networks. In Second, contribution authors have provided the relationship between two factors one among them is the video quality and the second one is maximum lifetime of WVSNs; which is very useful for the network design.

Energy aware sensor design

Energy aware sensor design is given by Yan in [33]. Author has proposed an energy-efficient strategy, which reduce energy consumption at sensor node level as well as at network level. They had designed sensor nodes that are energy aware and uses sleep/awake concepts. Sensor nodes estimate the distance between the transmitter and the receiver after that lowest transmission power needed is used. Furthermore, at network level energy saving also can be achieved by estimating the total energy consumption within the network under different network configurations and then by choosing the most energy-efficient network configuration.

2.3 Clustering

In 1997 author Lin has proposed cluster based architecture in WSN. This Clustered network architecture has three advantages one of which is bandwidth utilization at its maximum, second is bandwidth sharing and third advantage is to make network robust. In the proposed network architecture, nodes are organized into non overlapping clusters which are independently controlled [8]. It's a grouping of sensors in such a way that one sensor will become cluster head and others will be cluster members. Objective of clustering algorithm is to divide network into clusters to support multi hop communication and to support QoS.

Clustering algorithms are used to partition the network into a number of clusters. Cluster size is selected by considering two factors one is the tradeoff between spatial reuse of the channel and second criteria is delay minimization. Other constraints also considered like power consumption and geographical layout. Radio transmission power is also one factor to control cluster size but in this paper authors has worked and assumed that that transmission power is fixed and is same across the network. This paper has evaluated clustered network in terms of connectivity, transmission range and end to end throughput. The multi cluster architecture has been evaluated on MAISIE simulation environment. Simulation results showed d key tradeoffs between transmission ranges, and have shown the advantages of code separation and spatial reuse. Simulation shows that this clustered architecture provides a stable and efficient infrastructure for different types of traffic.

Further in [34] author said Clustering can contribute to more scalable behavior. Clustering is beneficial as it increases number of nodes, improved robustness, and more efficient resource utilization for many distributed sensor coordination tasks. Clustering comes under scalable network protocols. Clustering can also be taken as implementation of hierarchical networks [63, 67]. In clustered networks cluster heads take participation in election and also perform data aggregation. Authors have also proposed distributed clustering algorithm. Paper [72] has survey many clustering protocols. The clusters heads constructed based on nodes unique ID. The algorithm works as each node only broadcasts one cluster message. Before the algorithm stops within each formed cluster, nodes can communicate with each other in at most two hops. Clustering protocols sometime also taken as routing protocols [73]. Some routing protocols also use cluster base architecture for forwarding data [79]

First Clustering protocol was proposed by Heinzelman in 2000 [35] later modified in [36]. Authors targeted at issue of communication energy reduction, which give advantages of flexibility and energy efficiency, which give maximize network lifetime. Authors have also integrated the concept of "energy aware" such systems. The authors have proposed Low Energy Adaptive Clustering Hierarchy (LEACH) protocol. LEACH cluster based protocol uses local data fusion and classification which decrease the content of information that need to be transmitted to far located base station. In LEACH protocol, the sensor nodes which are cluster members use to transmit their data to a local cluster-head.

It works in two phases which are Setup Phase and Steady phase. LEACH works in distributive manner; it works on combination of two ideas one among them is energy-efficient cluster-based routing and the second is accessing media together with application-specific data aggregation to achieve improved system lifetime. The working of LEACH is divided into rounds. Every round has two phases one is a set-up phase in which the clusters are organized, followed by second phase named as steady-state phase in which aggregated data are transferred from the sensor nodes to their cluster head and on to the base station. It selects cluster heads on probability basis. At the starting level of LEACH, each member is assigned with the equal probability of being a cluster head. This probability is set as a function of a node's energy level relative to the aggregate energy remaining in the network. With Use of these probabilities, the nodes with higher energy are more likely to become cluster heads than nodes with less energy.

Once the nodes have elected themselves to be cluster heads each cluster head node broadcasts an advertisement message. After the end of single iteration, the winner cluster member is selected as cluster head. Each non-Cluster head node determines its cluster by choosing the Cluster Head that can be reached using the least communication energy. Second phase is steady-state, this phase is divided into frames. Senor nodes will send their sensory data to their cluster head at most one time per frame from allocated transmission slot. LEACH also ensures that no cluster head will be repetitively chased. This algorithm runs in iteration and in successive iteration cluster heads are reechoed. TT LEACH is given in [60].

LEACH protocol was further improved by author Lindsey in 2001 article published in next year 2002 [37]. Author has presented Power-efficient gathering in sensor information systems (PEGASIS) protocol which improved network life by factor 2 in contrast to LEACH protocol. PEGASIS has mainly focused on data aggregation and routing of data. It works on the basis of Travelling Salesman Problem (TSP). The working criteria of this protocol are chain based. Chain is formed among sensor nodes and each sensor node will communicate with only two nodes one for receiving from and one for transmitting to neighbors who are close. Collected data will be transmitted from node to next node of chain, processed to fuse, and finally a selected sensor node will send it to the base station. This approach will distribute the energy load evenly among the sensor nodes in the network. Here one more point that this protocol said is that the sensor nodes take turns transmitting to the base station so that the average energy spent by each node per round is reduced.

PEGASIS may use distributed approach to decide chain by greedy approach or alternatively centralized approach can be used in which base station will broadcast which chain will be used. The greedy approach works well for constructing the chain and it will be decided before the first round of communication. Chain construction start with furthest node from the BS the reason is nodes farther from the BS have close neighbors. Every sensor Node from chain performs data fusion except the end node. Sensor nodes perform fusion of received data from one side neighbor and their own. Fusion results a single packet of desired length and will transmit this packet to its second side neighbor. PEGASIS performs well by eliminating the overhead of dynamic cluster formation, secondly by minimizing the distance that non leader-nodes must transmit. PEGASIS also limits the number of transmissions and receives among all nodes, and using only one transmission to the BS per round. Turn to transmit the aggregated fused data towards base station is rotated among sensors to keep energy depletion balance. This rotation also gives benefit of robustness for randomly located sensor nodes. Simulation results have shown PEGASIS out performs LEACH as the size of Network increases.

LEACH-C was also proposed by Heinzelman in 2002 [36]. LEACH-C protocol is centralized in nature so works under high powered BS. It ensures that equal distribution will be for energy load. This protocol starts in manner that sensor nodes will send their current location and current residual energy to base station. Base station will calculate average energy of network on the basis of received energy information of individual sensor nodes. Those sensor nodes that are having residual energy less than the average network energy will not be considered for cluster heads role. Remaining sensors will be considered for clusters and base station will try to find k clusters among them using annealing algorithm. Annealing algorithm will try to select cluster members under cluster head by considering minimum energy required to communicate over distance between cluster heads and their cluster members. Once BS is done with decision about cluster heads and their cluster members, BS will broadcast the list of selected Cluster heads. List also contains ID of cluster heads. When this list received at nodes under that region then only cluster heads those are having ID from that list will remain in active state rest all nodes go to sleep stage. Sleep nodes will keep waiting for TDMA schedule and will transmit sensory inputs at the end. LEACH-C has good performance as it will produce better clusters in contrast of distributive clustering and also disperse cluster heads throughout the sensor network in better way.

In 2001 [38] TEEN (Threshold sensitive Energy-Efficient sensor Network) a distributive power base protocol was proposed. This paper has separated two types of network on basis of their functioning one is proactive and other is reactive. TEEN protocol works for reactive type networks. Whenever there are some variations in threshold value, the sensors nodes would transmit instead of their regular practice of sensing these types of networks are called reactive networks. Protocol starts working as like at every cluster change time, the cluster-head broadcasts to its members two values; one is a hard threshold and other is a soft threshold and attributes as well. The sensor nodes keep sensing data and store this data to a variable called SV (Sensed Value). When both conditions are satisfied then sensor nodes will start transmission during present transmission cycle. Out of two conditions one is when the current value of the sensed attribute is greater than the hard threshold, and the second is when the current value of the sensed attribute differs from SV by the amount equal to or greater than the soft threshold. Whenever a sensor node has transmitted data, the variable SV is set equal to the current value of the sensed attribute. The aim of these two thresholds is to reduce number of transmissions. Hard threshold limits the transmissions by posing condition according to which node will transmit once the sensory values posses range of interest. On the other hand soft threshold limit the transmissions by posing condition that nodes will transmit when there is small change identified in hard threshold. This protocol is benefited with advantage of instant notification to user in case of data with time sensitivity. This protocol is best suitable for real time applications which are having time sensitivity, for example, all applications where change point detection is required. TEEN protocol shows better accuracy when soft threshold is set to a smaller value, but yes at the expense of increased energy consumption. TEEN protocol have the main drawback which is if the thresholds are not reached, means no attribute can satisfy both condition in that case the sensor nodes will never communicate, the user will not get any data from the network at all and also will not come to know even if all the nodes die. So this is also concluded that this protocol is not good for regular data gathering applications. APTEEN protocol has covered some drawbacks of TEEN protocol [75] One more Weighted Clustering Algorithm(WCA) was introduced in 2002 that was also for reactive networks but it was considered only for Adhoc networks, sensor network's futures and requirements was not considered [76].

Further in 2004 author Younis has proposed Hybrid Energy-Efficient Distributed clustering (HEED) in paper [10]. HEED protocol is again distributive in nature and also power base protocol. Another distinguish feature of HEED is that it's a multilevel clustering protocol which also support multi hop communication. HEED protocol works to select cluster heads among sensor nodes on the basis of two factors one is node residual energy and the second factor is communication cost between intra-cluster. HEED has an advantage that it does not require special node capabilities, such as in HEED, location-awareness, does not depend on node distribution *etc.* In HEED, each node is mapped to exactly one cluster and can directly communicate with its CH. Cluster head selection in HEED protocol is primarily based on the remaining energy of each sensor node. Remaining energy is estimated current energy in the node. Secondary clustering parameter that authors has considered is intra-cluster communication cost. The algorithm include three phases: first phase is initialization phase, second phase is named as repletion phase and the last phase is known as finalizing phase. In first phase broadcast cost is calculated according to nodes under range of communication and also cluster head probability for each sensor node is calculated. Second phase is repeat phase; this phase decides tentative cluster heads. If any node is having probability equal to one then this sensor node will be declared as final cluster heads otherwise least cost head is taken as tentative cluster head. Cluster heads are finalized in last phase. And sensor nodes become cluster member under least cost reachable cluster head. HEED protocol is benefited with parameters such as the minimum selection probability and network operation interval due to which HEED protocol easily tuned to optimize resource usage according to the network density and application requirements. HEED protocol also ensure that it will terminates in a constant number of iterations, independent of network diameter.

In 2005 author Li has proposed one more distributive clustering protocol which is named as energy-efficient unequal clustering protocol (EEUC) in [39]. EEUC is a protocol, where residual energy is the base for selection of cluster head. EEUC protocol has targeted at hot spot problem and early network partition problem. It is stated that this protocol can help to avoid generation of energy holes within networks. The problem in tradition cluster schemes was that cluster heads near the base station get drained their energy early as contrast to cluster heads far from base station. The reason is obvious because near located cluster heads have extra overhead because of intercluster communication. Far located cluster heads are not able to transmit their data to base station with single hop communication method so they opt multi hop communication method. Multihop communication method forces near located cluster heads to support intercluster data from far located cluster heads.

EEUC works differently as it divides the nodes into clusters of unequal size of sensors. Cluster partition is done on the basis of distance from the base station. The clusters closer to the base station will be of smaller sizes so that they consume less energy in intracluster communication. And this energy can be negotiated by data forwarding from cluster heads those are farther away from the base station. This protocol has showed good performance in terms of throughput. Algorithm starts working by broadcasting hello message by which distance of each node from base station is calculated. Next step calculates probability of sensor node on threshold value basis. In iteration nodes with same probability under a range of communication will be elected as tentative cluster heads. Rest all nodes go to sleep mode at this stage. Further procedure continue with selection of final cluster heads from set of elected tentative cluster heads, and this final cluster heads will consider distance from base station. It is evaluated that in EEUC cluster heads consumes less energy in contrast of LEACH, and is about the same as that in HEED. EEUC supports well balanced energy consumption among nodes. Authors had shown results that EEUC protocol has targeted hot spot problem by reducing time intervals between time at first node dies and the time when last node dies.

In 2006 one clustering protocol came in research name as Power Aware Dynamic Clustering Protocol (PADCP) in paper [40]. PADCP is low energy protocol and different from other clustering protocols as it adapt many adaptive schemes. These adaptive schemes include Dynamic Cluster range (DCR), second is Dynamic Transmission Power (DTR) and also it support Cluster Heads re-election. DCR means some cluster head low range if they have small size or according to large cluster size it may have more range; this factor also give benefits that PADCP can support variable sized clusters. DTR is useful for communication in between cluster heads. Cluster heads can set their transmission range to forward or receive data from neighbor cluster heads. Cluster Heads re-election is very beneficial feature to keep clusters active all time and load management within cluster. It works when active cluster head get down due to energy fall up to a threshold value then cluster member re elect their cluster head which will having maximum residual energy. PADCP focus on topology control as it assumes non stationary nodes as well. It supports different power level which can also control energy consumption. PADCP has shown improvement in network lifetime also can say modified HEED protocol.

In [41] Power-efficient and adaptive clustering hierarchy protocol (PEACH) has been introduced. PEACH protocol works with both types of networks either location-unaware (LU) or can for location-aware (LA) wireless sensor networks. It operates on probabilistic energy-aware routing and removes the difficulty of cluster formation.

PEACH protocol has targeted on energy consumed during cluster formation. Author surveyed and concluded most clustering protocols spend their large amount of energy and time for cluster formation. In this paper authors have aimed at saving this energy as well. This protocol is a multileveled protocol works to form adaptive clusters according to overheard information. PEGASIS has used the greedy TSP (Travelling salesman problem) algorithm to form a chain to connect every sensor node. From this chain nodes are restricted to communicates with two of their close neighbors. Simulation results has shown that the performance of LUPEACH and LAPEACH protocols provides multiple benefits firstly long network lifetime, secondly more scalability and lastly less and accurate energy consumption, and greater scalability than compared with existing clustering protocols. It is also very important point that Delay is sometime not acceptable especially in real time applications.

Author Yang has invented a centralized protocol named as SHORT in [42]**.** SHORT is more than clustering also focuses on routing of data to reduce delay. It's a centralized chain based protocol in which clusters have to be selected by base station that is reason why it requires high-power base station. Protocol works by effectively generating simultaneous communication pairs and identifying the shortest hop. The author has proved by experimental evaluation that SHORT protocol results better performance in terms of delay and energy in contrast to present chaining aggregation clustering protocols. But this is well understood fact that centralized protocol is having limitation of high power base station requirement.

One more distributive protocol was in work which was proposed by author Lung in 2010. DHAC is a distributive hierarchical multilevel agglomerative clustering proposed in [43]. It is actually an extension to work done on HAC. HAC protocol is different from other clustering protocols as it works in Bottom up approach in contrast all other works on TOP- Down approach. HAC protocol uses global knowledge about network and also uses some already defined methods. Protocols first works on the upper level and select some sensor nodes as CHs. Cluster members are assigned under clusters in next level usually this level is taken as lower level. The author in this paper has been simple in defining six-step clustering in which the work is done by making group of similar types of nodes together with the help of a resemblance matrix. In DHAC cluster formation needs information about one hop neighbors. DHAC avoids re-clustering and attain consistent energy indulgence through the entire network. DHAC have three different features First DHAC works on bottom-up approach. It starts clustering at lower level by sensors working with one-hop neighbor information. Second feature of DHAC is that it can work on different input data types. Third, DHAC gives energy efficiency at max and also it performs clustering single time during initial stage. DHAC perform cluster formation in six steps; among which first is collecting neighboring information, second is design of resemble matrix, third step is execution of DHAC algorithm, forth step is about cutting hierarchical tree followed by fifth step is about controlling the minimum cluster size in final sixth step cluster heads are chased. As DHAC perform clustering only once then cluster heads are maintained by automatic cluster heads rotation. Author also showed its implementation on NS2 simulator and it is proved that DHAC perform in better way during light traffic or less load in the network. Authors had shown DHAC has defeated both of LEACH protocol and LEACH-C protocol on performance basis.

In paper [44] one more distributed Energy Efficient Cluster formation (EECF) protocol was proposed. EECF protocol form cluster and elect cluster heads on three-way message exchange. Three way message exchanges occur in between sensor nodes those are regular nodes, Cluster Heads and one hope neighbor. Cluster Heads are elected on basis of respective residual energies and degrees. One more point that EECF include is score advertisement broadcast (SAB). Each sensor node broadcast its SAB to neighbors. SAB value is also get priority when some sensor node has to get promoted as cluster head. Once any node is promoted as cluster head all near by sensors get attached to it as cluster members. Sensor nodes will wait for their neighbors' decision before final decision of joining which cluster head they will join as cluster members. Cluster heads also act as relay they get data from peer cluster heads. Cluster heads form a route up to sink node; they forward data to next cluster head peer of route. Cluster heads also compare SAB for forwarding data to next peer cluster head. EECF is aimed at improving network lifetime; one factor more can affect performance of EECF is distribution. Different energy distribution can vary performance of this protocol. EECF is benefited with message exchange complexity of O (1). In worst case convergence time complexity is O (N)

In 2014 the author Tarhani has targeted on one more issue of relaying in clustering [45]. Authors have tried to solve the problem which cluster heads faces. As it's a fact that cluster heads also act as relay for neighbor cluster heads, they needs to forward data that comes from other clusters heads to satisfy multi hop communication chain. Due to which they need to spend their more energy in forwarding instead they should preserve their energy for collecting and aggregating to sensory information received from cluster members. That's why SEECH protocol has proposed new approach for selecting cluster heads and relay nodes separately. Cluster heads will not spend their energy on relay services. Only relay nodes will perform forwarding tasks. Cluster head's energy will be saved. SEECH protocol also proposed a new distance-based algorithm. Simulation is carried for different parameter on networks and overall SEECH has shown that network lifetime in SEECH /is better in contrast of LEACH. We are presenting survey of clustering protocols in tabular manner in Table 2.1.

2.4 Change Point Detection

Change point detection is has role in many real-life applications. Name of such real time based application include environmental monitoring, Surveillance monitoring, Habitat monitoring, Industrial monitoring, Target tracking and Health monitoring in medical fields etc. Change point detection also has application in security areas. In WSN one more term 'Outlier detection' is also used in place of change point detection. Term Outlier is first introduced in paper [46]. The author defined outlier is inconsistent observation in contrast of other observations belong from same set [90].

Later term outlier was also discussed for statistical data in [47]. Author defined an outlier is an observation, which deviates so much from other observations as to arouse suspicions that it was generated by a different mechanism. Later on change point detection i.e. outlier has become important in multiple fields [71]. Change point topic was overviewed with TEEN protocol according to which sensors will react only when an event occurs or when change is above to threshold values. Here we will present our survey related to change point detection on methods bases. Change point detection can be done via different methods as shown in figure 2.3.

Table 2.1- Clustering protocols with Feature survey

Clustering Protocol	Centralized	Distributive	Multi hope Inter cluster communication	Location Awareness	Power Base	Multilevel Clustering
LEACH		√				
HEED		√	√		√	√
LEACH-C	√			√	√	
TEEN		√			√	
PEACH		√		√		
SHORT	√					
EEUC		√	√			
DHAC		√				√
EECF		√	√		√	
PADCP		√	√		√	
SEECH		√			√	

Fuzzy logic

Fuzzy logic is very useful method in change point detection. The process starts with vague data; data is fuzzyfied in next step is to build fuzzy sets also probability functions are used then after results are concluded from fuzzy rules. In 1995 the author Dexter has proposed how fuzzy theory is used in fault detection [91].

In 1999 author Metternicht has used same approach for computation of changes occurred in remotely sensed area for map revision [65] The author has formulated fuzzy sets and fuzzy rules for defining the likelihood of changes detected from remotely sensed data then after they used decision support system to conclude change. Authors had shown that fuzzy logic has given better change detection results in contrast of traditional systems like image ratio etc. results are shown to prove how change point detection using fuzzy theory help to conclude map revision more accurately.

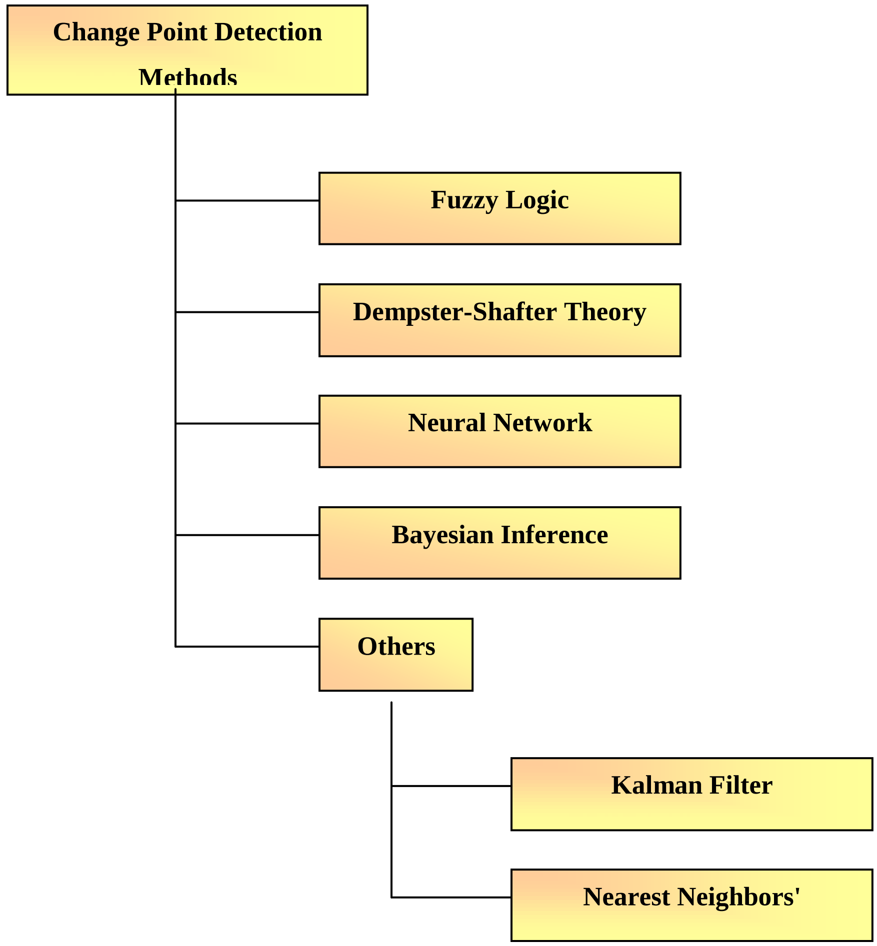


Figure 2.3- Different methods for change point detection

In paper [48] author had worked on application concerned on tracking and denoising of image sequence. The work has contributed fuzzy recursive motion detector. The author has worked by inputting noisy sequence with fuzzy logic motion detector in order to determine the degree of motion confidence. They proposed motion detection algorithm; concentrating on two criteria one is robustness to noise and second is changing illumination conditions and motion blur in temporal recursive denoising.

Further in 2011 the author D'Angelo used the method of fuzzy logic to find faults in machines example of such machine is; induction machine stator-winding [49]. The author has used the approach of change point detection for improved forecast of the new failure detection procedure against false alarms, combined with a good sensitivity that allows the detection of rather small fault signals. They had defined and implemented two steps procedure. First step is transformation of initial data using fuzzy clustering. Second is change point detection with the help of Metropolis–Hastings algorithm.

Fuzzy logic can also be merged with genetic algorithms for fault detection as done by author Jack in [95]

Neural network

Neural networks also work for anomaly or fault detection. In [50] the author had worked for change point detection in inputs of sensory motes. Authors have given and executed two new neural network based algorithms named as one is ART and other is named as Fuzzy ART. These two algorithms work as classifier to categories sensory data; this categorization is performed without any supervision. Whenever there is some variation in data inputs then alarm is triggered.

In [51] research has been carried performed for fault detection in WSN using neural networks. This Work is done for real-time implementation of WSN worked on sensors operating system TinyOS. The author has used the concept of recurrent neural networks (RNNs) to represent the sensor node. According to which every sensor node will receive data from its neighboring sensor nodes. Dynamic RNNs is constituted by a number of dynamic nodes and these nodes provide feedback to there input with confidence factor, after that RNN output for sensor node will be calculated. Error is calculated by the difference of predicted value and actual physical measurements. If this error is larger than a predefined threshold, then it indicated there is a fault in some sensor.

Further in 2010 author Siripanadom has shown use of neural network for anomaly detection in WSN [52]. The method used for change detection in this paper is based on learning neural network. The method is named as self-organizing map (SOM) which uses DWT for reducing data as well. The solution they have presented here is having advantages of limited storage and computing costs. However SOM method requires processing time which increases with the size of input data that’s why author had used DWT to reduce the input data size, without losing the significant features of the data for anomaly detection. SOM uses statistic from input data which further encoded in to weights with self learning. A new observation data set can be considered abnormal if the distance between the weight vector of the winning neuron and the new state vector, given by new state vector is greater than a certain percentage.

Dempster-Shafter Theory (DST)

The author Zadeh in 1986 has presented a view of the Dempster-Shafer Theory of Evidence and also it's Implication for the Rule of Combination [53]. The Dempster-Shafer theory works based on two ideas first among them is the obtaining degrees of belief for one question from subjective probabilities for a related question, and second is use of Dempster's rule for combining such degrees of belief when they are based on independent items of evidence. Dempster-Shafer theory of evidence has attracted considerable attention as a promising method of dealing with some of the basic problems arising in combination of evidence and data fusion [53]. DST produces a judgment value between 0 and 1 that reflects the degree of belief in that judgment.

It has been shown in [54] that DST is more suitable to use for structural health monitoring in contrast of conditional probability. The author has used the concept of hierarchical WSN and proposed a method named as structural health monitoring. Sensor nodes send the sensory values to their cluster heads and after that cluster heads process and aggregate these values and make a local report. The Judgment about health status will follow the majority where a correct majority decision requires.

Further in 2012 in paper [55] DST is used for multi sensor data fusion. The aim of this theory is to detect change in dynamic circumstances. The authors have presented a inference method worked for swift context inference. DST is merging the belief plausibility and uncertainty all together, after that result has been compared to the highest belief and the lowest uncertainty of all focal elements to find the cause of the change.

DST theory has been also used for change detection of mechanical devices like pumps etc. The author Guo has showed its use in multi variable system for fault diagnosis [92]. DST has also proved its efficiency in detecting damages sites which are compound structural damaged [93]. Recently DST has been used for fault detection in railway system by the author Wei [94]

Bayesian Inference Theory

Bayesian theory is used where conditional probability is needed to implement. Bayesian theory is a method of statistical inference which works on basis of Baye’s theorem. According to Bayes theorem

$$P(H \mid E) = \frac{P(E \mid H) \cdot P(H)}{P(E)}$$

Where H is used to denote hypothesis
E is used to denote new data
P(H) used is for prior probability
P(E) is represented for model evidence
P(E/H) is probability of evidence E when H hypothesis is given
P(H/E) is reflected for posterior probability

Bayesian inference is closely related to Bayesian probability when it comes to decision probability. Bayesian start with taking values of prior distribution then describes the degree of belief. Bayesian belief function is useful in outlier detection [72]. In paper [66] authors has used Bayesian decision theory for change point detection in remotely sensed images. Technique used by author is based on unsupervised estimation. Aim of this paper was to find recognition of changes automatically in images. Images are raw images of same area taken at different point of time. Three algorithms are proposed by authors one is Reduced Parzen estimation (RPE), second is Expectation maximization (EM) and third one is Markov Random Field (MRF). Bayesian theory also posses’ great importance in the field of change point detection for images.

In [56] work is done for change point detection targeted on synthetic aperture radar (SAR) images. Method used by authors is Bayesian inference implemented along markov chain Monte Carlo (MCMC). Authors have studied and worked with two estimators one is minimum mean square error and second estimator is maximizing a posteriori. Authors have also clearly stated that in practical these two simulators cannot be used simultaneously. MCMC methods are used to simulate the posterior distribution of the change point positions. The work is aimed at edge detection in SAR images offline way. Offline edge detection is performed line-by-line and column-by-column.

In [57] work has been carried out with naïve Bayesian classifier for WSN security. This agent based classifier is set up in-between cluster head and base station. The author also had shown with experimental results on NS2 that how this classifier is successful for detection of abnormal events in WSN security. In paper [58] rule based classifier which is based on naïve Bayesian algorithm is used to detect outlier in WSN security. This paper proposed a algorithm for training n and test data set for detection as well as rectification of outlier and author had simulated this work on NS2. The author has explained how the essence of the Bayesian approach for providing a mathematical rule also explaining how you should change your existing beliefs in the light of new evidence.

2.5 Gaps in Study

- Out of many energy aware clustering schemes none has used updated energies of sensors like in LEACH remaining energy had been used to find probability only in setup phase.
- Some clustering protocols are taking Cost of communication but like in HEED Snbr is considered once in initialize phase.
- No combined application oriented clustering protocol has been introduced. Only TEEN protocol has combined clustering with change point detection; but this was only for reactive networks.

CHAPTER 3 MUTUAL EXCLUSIVE DISTRIBUTIVE CLUSTERING

3.1 Introduction

Clustering protocols are needed in WSN for making them more sustainable, scalable and efficient. Number of clustering protocols was discussed in [9]. After intense survey it is found that distributive clustering protocols are easy to develop and no extra setup cost is required so distributive protocols must get preference although their disadvantage is that they are not always reliable. We in this chapter will present our proposed protocol named as Mutual Exclusive Distributive Clustering (MEDC). Section 3.2 is about background for this work. Section 3.3.1 will discuss some network model assumptions related to our work. 3.3.2 Will discuss Radio model of communication for protocol. Section 3.2.3 will present gist of this protocol and also enumerate an example to show working of MEDC. Section 3.4 will be about algorithm which is diagrammatically represented as flow chart in section 3.5. Section 3.6 is about experimental evaluation. We have simulated MEDC protocol on MATLAB. Results are taken on varying two parameters among them one is number of sensor nodes and second is range of communication. For simulation we had taken two values for sensors one is 100 another is 200. For Range of communication we had simulated on four values 20, 40, 60 and 80. Section 3.7 about result discussion after that work is concluded in section 3.8.

3.2 Background

Clustering protocols select cluster heads to reduce communication cost so that network can live to maximum. A number of clustering protocols has been produced which may have different working parameters and features. We had taken residual energy or can say power based clustering protocols. HEED is one of them we will discuss HEED protocol in detail. We had considered this protocol as base protocol of our work and compared performance with it. The reason for considering HEED is our network assumptions. HEED protocol has worked on similar network assumptions [10]. HEED is distributive in nature and power base protocol. It selects cluster heads on basis of two factors one is nodes remaining energy and second factor is intra-cluster communication cost. HEED protocol does need extra capable sensor nodes for location awareness even this protocol need not any special sensor node distribution methods. Cluster head selection in HEED protocol is primarily based on the remaining energy of each sensor node. Remaining energy is estimated current energy in the node. Secondary clustering parameter that authors has considered is intra-cluster communication cost. This protocol includes three phases. First phase is known as Initialization phase, in which broadcast cost is calculated according to nodes under range of communication and also cluster head probability for each sensor node is calculated. Second phase is repeat phase; this phase decides tentative cluster heads. If any node is having probability equal to one then this sensor node will be declared as final cluster heads otherwise least cost head is taken as tentative cluster head. Last phase is called finalizing phase; Cluster heads are finalized in this phase and sensor nodes become cluster member under least cost reachable cluster head. HEED protocol works with minimum selection probability and use minimum network operation interval due to which HEED protocol easily tuned to optimize resource usage according to the network density and application requirements. HEED protocol also ensure that it will terminates in a constant number of iterations, independent of network diameter [10].

3.3 MEDC

3.3.1 Network Model assumption

- Sensor nodes are of homogeneous type means they have the same capabilities and resources like battery power etc.
- Sensor nodes are stationary deployed in local region for monitoring and the data sink is located far from the sensing field.
- Network is formed of location unaware sensor nodes; nodes does not having any capability like GPS
- Sensory data is aggregated at different levels and sent to data sink generally called base station at regular period of time
- A different identifier will be used for each sensor node.
- Communication is done on symmetric links. Communication can be bidirectional.

3.3.2 Radio model Equations

Sensors energy is dissipated on transmission and receiving activity along with sensing. If any sensor want to transmit t bits to a node located at distance d. Then energy dissipation is calculated by equation 3.1

- Transmission Energy Dissipation
    - Etx (t) =t* Eelec. + Etx_amp (t, d) (3.1)
    - Etx_amp (t, d) = t*d^2*Efs (3.2)
    - Etx_amp (t, d) = t*d^4*Emp (3.3)

Where Etx represent transmission energy will be calculated from electron energy (Eelec) and amplification energy (Etx_amp). Amplification energy consumption also varies depending on free space communication or multipath communication equations 3.2 and 3.3 are shown. Free space energy and multipath energy consumption is represented as Efs and Emp respectively. If any sensor node is receiving t bit from any sensor node then energy dissipated for t bits is represented as Erx

- Receiving Energy Dissipation
    - Erx (t) = t* Eelec (3.4)

3.3.3 Gist of MEDC

Mutual Exclusive Distributive Clustering (MEDC) will form clusters of sensor on basis of mutual exclusion algorithm. MEDC Clustering protocol chooses cluster heads on basis of mutual exclusion algorithm from the number of sensors. Within a range of communication (says Rc), only that sensor node which is having maximum of residue energy will be chased as cluster head from that range. This protocol is having four steps in single iteration and each iteration will decide new cluster heads which are having maximum residue energy under a range of communication. Under Rc, the Cluster head sensor node will be decided on factor of remaining energy. Sensor node which one is having highest remaining energy among the sensors node under Rc will be chosen as cluster head. The idea of Mutual exclusion works by message passing system we will say it by advertisements. Why we are considering the remaining energy factor , is because sensor's remaining energy will be different for each sensor nodes infect it depend on how much each sensor node spends energy for sending and receiving advertisements up to last iteration. Which in turn how much advertisement depends on how many sensors were under range of communication (Rc)

Here we want to highlight why we are considering only residual energy for choosing cluster heads not distance factor. Some reader may say distance can also be good factor for selection. But here are considering sensors are deployed in local region say in 500* 500 meters; base station is far away say at 25 Km. So relative distance for all sensor nodes will be same for all sensors. As shown in diagram for two sensor nodes.

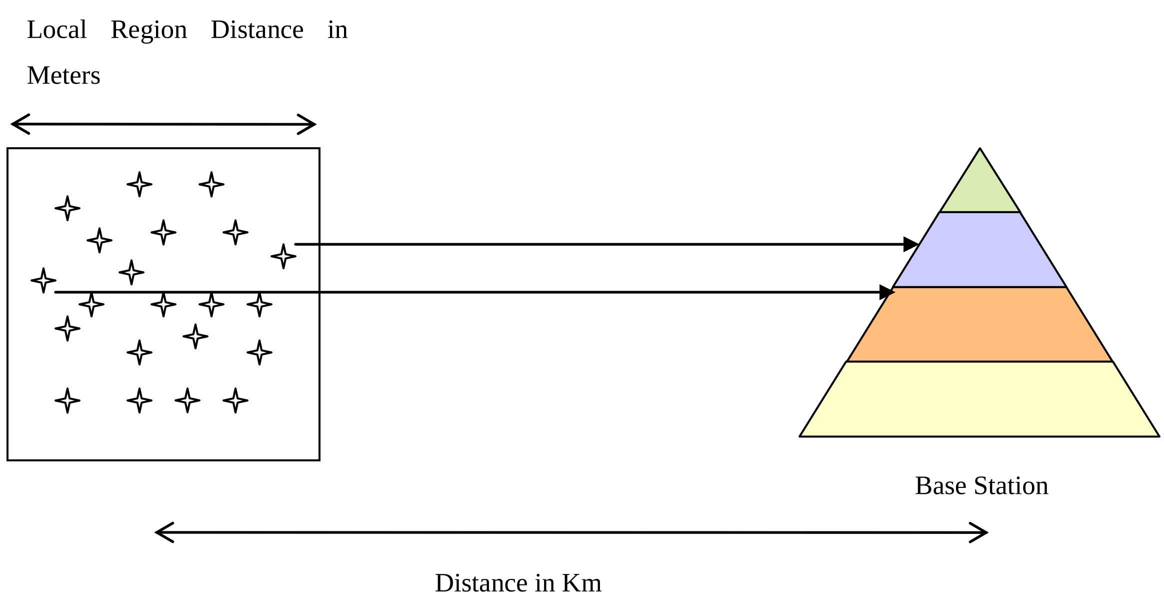


Figure 3.1-Scenario of Deployment

Secondly here we consider communication is multipath so in between relay nodes will be there so relay nodes will forward the data without consideration from which sensor they are receiving data for forwarding. The third point is we cannot burden near located nodes for relay service in that case they deplete their energy and cause portioning of network problem.

Steps of MEDC are as follows:

Step 1

Sensors will send advertisement packets to all other nodes that come under their range of communication say R_c. Advertisement packets consists residual (remaining) energy of sensors nodes. Each sensor node will precipitate in sending and receiving of advertisements. Received advertisements will be saved in queues of sensor nodes

Step 2

All sensors nodes will check their queue. Each sensor node will send OK message after selecting sensors from queue. The selection criteria for OK message is; comparison of residual energies. Sensor node will select other sensors from queue which are having more residual energy then its own. The sensor node which has sent any OK message to other sensor from queue that cannot be cluster head for this iteration. If there exists any node that is not sending any OK message then it will wait.

Step 3

Step 3 will decide which sensors will act as cluster heads and which will be cluster members. As CDMA slot ends sensors will check their own status. The sensor node who has not sent any OK message who has not found any node from queue with more energy then owns it means this node will declare itself as cluster head. Cluster head declaration message will be sent to all nodes that come under its range of communication.

Step 4

Cluster head sensor nodes will collect and aggregate data from cluster members under $R_{c.}$ Cluster heads will transmit the aggregated data to sink node. Cluster head rotation will be done in next iteration. Next iteration starts from step 1.

Example

Here we want to show how MEDC will run and select cluster heads. In following figure five sensors are randomly deployed and they are identified as IDi. When they are deployed they were having same energy say 0.2 see figure 3.2. Figure 3.2 shows that ID1, ID2, ID3, ID4, ID5 are the sensors within range Rc. MEDC protocol will start working. In step 1 sensor nodes start sending and receiving advertisements under range of communication, Sensor nodes deplete their energy on sending and receiving advertisements due to which it may possible that nodes have different residual energies refer figure 3.3. Sensor nodes queue incoming advertisements here. Sensor nodes ID1, ID2, ID3, ID4 and ID5 will advertise their remaining battery life 0.12, 0.14, 0.19, 0.13

and 0.11 respectively. Each sensor node will maintain queue for incoming advertisements. Sensor queues are shown in table 3.1.

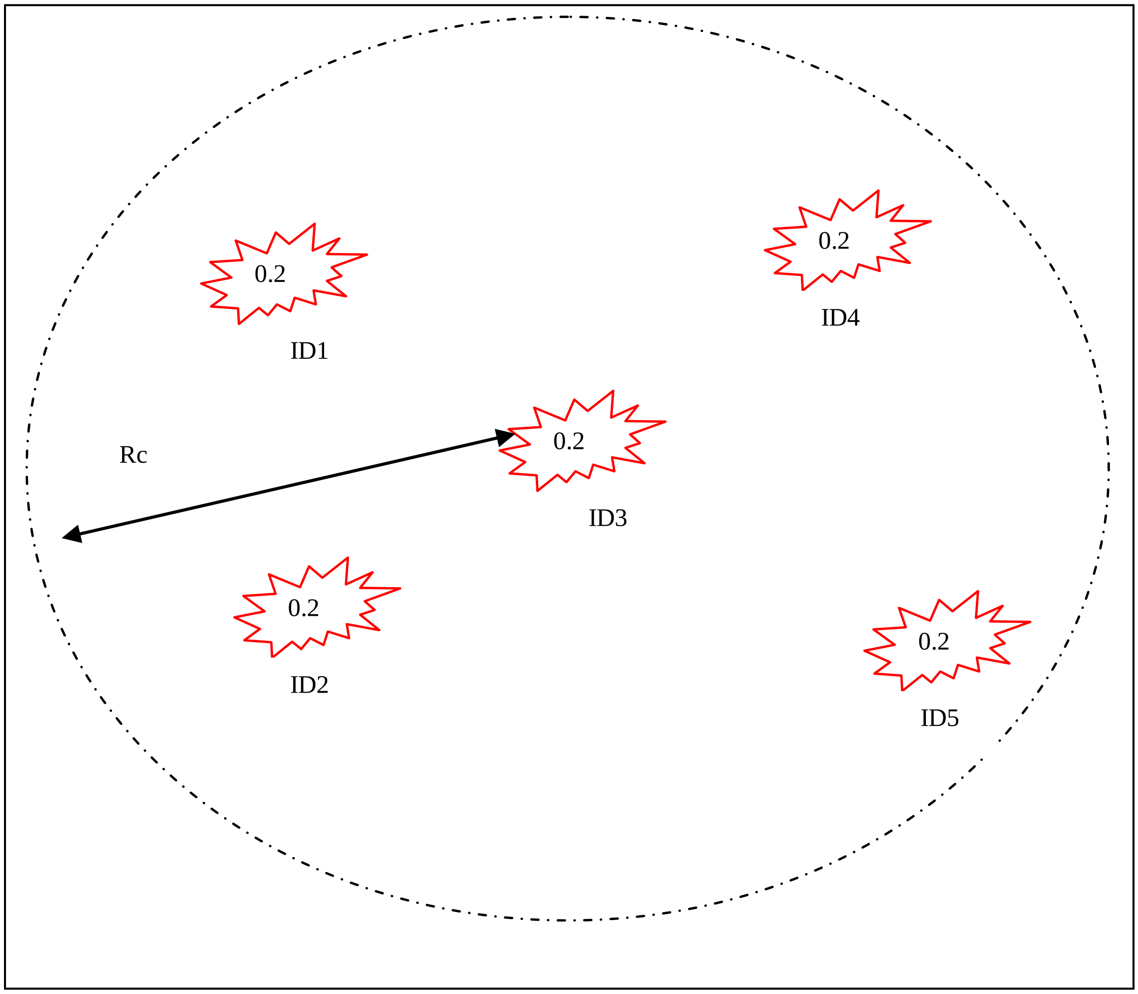


Figure 3.2- Initial Deployment of sensor

Step 2 of algorithm states that each sensor node will send OK message to sensor nodes from queue those are having more residual energy in contrast of their own residual energy see figure 3.3. Table 3.2 has shown that each node has sent OK message to which nodes. For example sensor node ID1 has received advertisement from ID2, ID3, ID4 and ID5. But from queue it will send OK message to ID2, ID3 and ID4; but not to ID5. The reason is ID5 is having residual energy 0.11 which is less than residual energy of its own i.e. 0.12 so ID1 will not send OK message to ID5. Same step will be executed for each sensor.

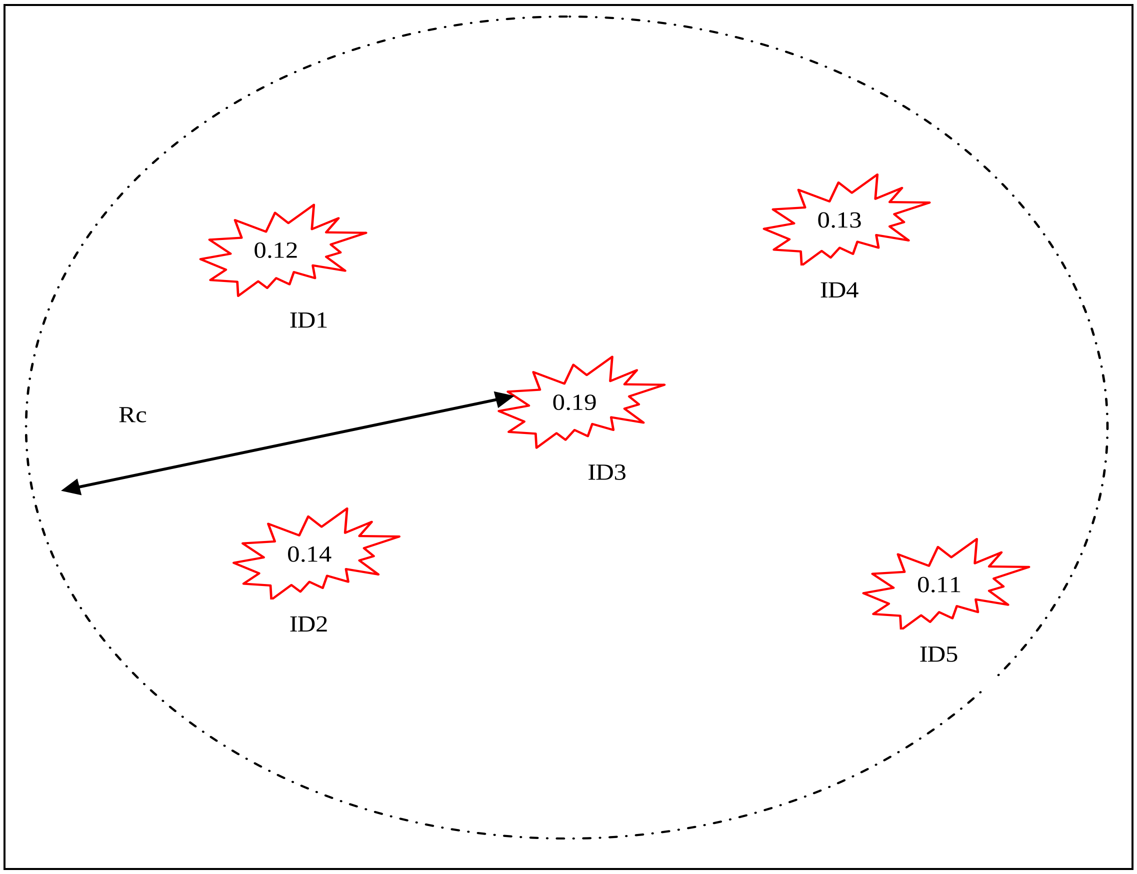


Figure 3.3- Sensor nodes after sending and receiving advertisements

In step 3 sensor nodes will check their status. If any sensor who has not sent any OK message this means that this node is having maximum energy under range of communication.

Here from table we came to know that node ID3 has not sent any OK message. Here sensor node ID3 will declare it self as cluster head and send this declaration message to ID1, ID2, ID4 and ID5 from range of communication see figure 3.4 mentioned as green colored.

During Step 4 cluster head ID3 will receive sensor data from cluster members; aggregate it and transmit it to base station. Next iteration will start after TDMA slot which is required to transfer data up to base station. Cluster heads will be change in successive iteration. Iteration will start from step1 to step 4.

Table 3.1- Sensor Queue

Nodes	Queues
ID1	ID2 (0.14), ID3 (0.19), ID4 (0.13), ID5 (0.11)
ID2	ID1 (0.12), ID3 (0.19), ID4 (0.13), ID5 (0.11)
ID3	ID1 (0.12), ID2 (0.14), ID4 (0.13), ID5 (0.11)
ID4	ID1 (0.12), ID2 (0.14), ID3 (0.19), ID5 (0.11)
ID5	ID1 (0.12), ID2 (0.14), ID3 (0.19), ID4 (0.13)

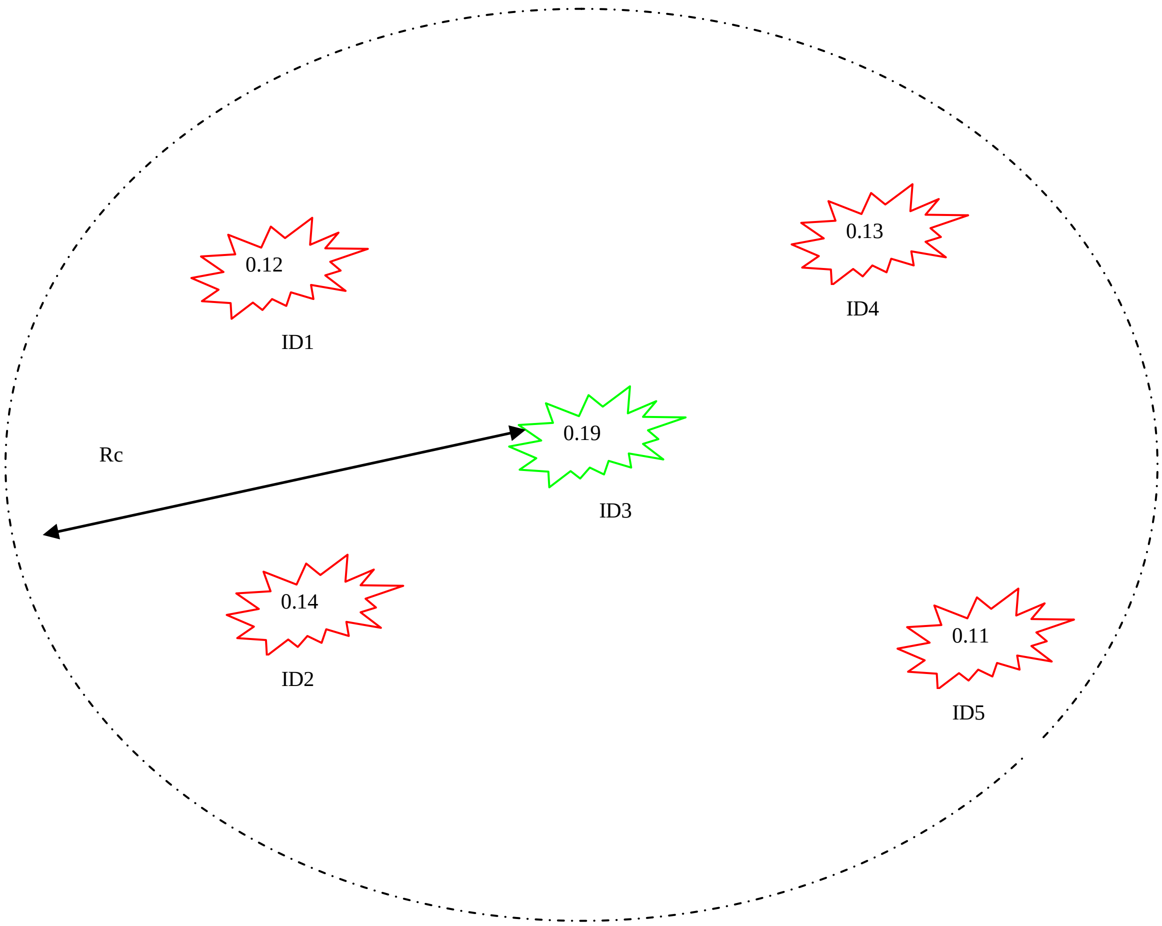


Figure 3.4 -Node ID3 will become Cluster Head

Table 3.2 Sent OK messages

Nodes	Sent OK message to
ID1	ID2, ID3, ID4
ID2	ID3
ID3	Nil
ID4	ID2, ID3
ID5	ID1, ID2, ID3, ID4

It might possible that some issue can occur out of which we are listed some of them and Issue resolving in MEDC

1. If a sensor gets more than one cluster head's cluster_head_declaration.

   Solution to this issue can be that the sensor will consider its cluster head with lower ID from many cluster head's and it will sensed it sensory values to that cluster head only.

2. If a sensor does not get any cluster_head_declaration.

   Solution to this issue can be that the sensor waits till the next iteration.

3. If a sensor does not get any OK message.

   Solution to this issue can be that the sensor will act as cluster head for its own.

4. If a sensor has sent cluster_head_declaration and it get some other sensors cluster_head_declaration

   Solution to this will be on basis of timestamp ordering. It might possible that message has been transmitted in previous iteration and get received in next iteration may be due to congestion. Message with lower timestamp will be ignored. Events are logically synchronized each event carry timestamp along it.

3.4 Algorithm MEDC

IDi : ID of node i.

R_f : Radius of frequency/ Range of communication.

Qi: Queue of sensor i.

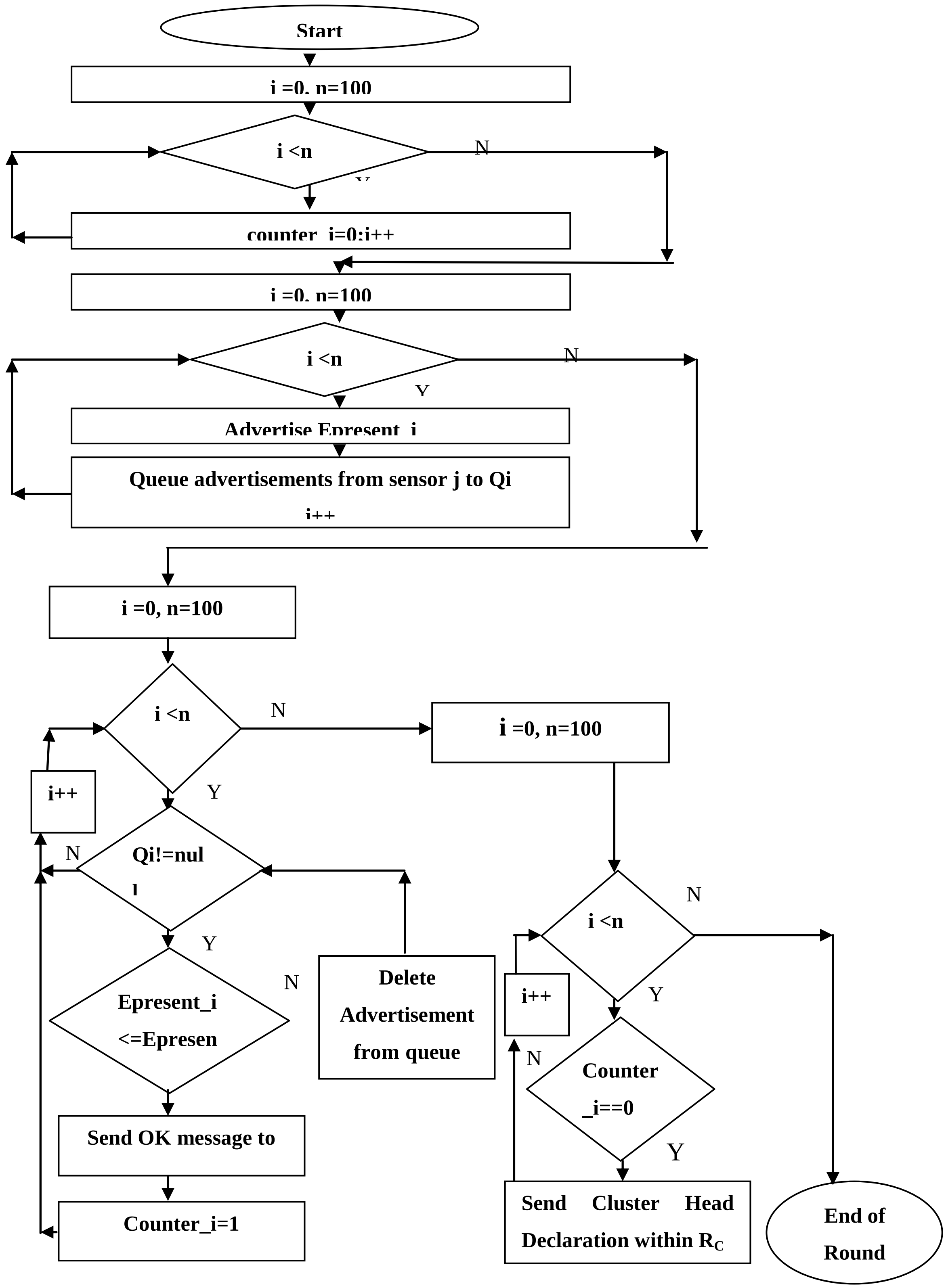
Epresent_i: Present energy of node i.

Procedure Cluster_formation (n)

```
1.      For each Next iteration
2.      For each IDi
3.              Counter =0
4.      For each IDj within Rf of IDi
5               Advertise Epresent_i
6.      For each IDi
7.              Put all incoming advertisements from sensors j into Qi
8.      For each IDi
9.                While Qi is not empty
10.                       If Epresent_i <= Epresent_j
11.                               Send ok message to IDj
12.                               Counter = 1
13.                               Else
14.                               Delete this advertisement from queue
15.      For each IDi
16.                     If counter =0
17.                     Send cluster_head_declaration message to IDj within Rf
```

Procedure cluster formation will be executed on each sensor with start of TDMA slot. One TDMA slot has setup phase in which clusters are formed rest part is steady phase cluster heads collect information from cluster members. Aggregation is done at cluster heads and then this aggregated information is sent to base station at the end of TDMA slot. TDMA slots are usually taken by considering maximum time transmission can take.

3.5 Flow Chart



3.6 Experimental Evaluation

For simulation of MEDC on MATLAB we had taken following parameters. Shown In table 3.3.

Table 3.3- Simulation parameters for MEDC

Parameters	**Abbreviation**	**Values**	**Units**
Random field x axis	Xm	100	Meters
Random field y axis	Ym	100	Meters
Initial energy of sensor	Eo	0.05	Joule
Total number of sensor	N	100,200	
Transmission energy	Etx	50*1.E-12	Joule per bit
Receiving energy per bit	Erx	50*1.E-12	Joule per bit
Free space energy per bit	Efs	10*1.E-12	Joule per bit
Data aggregation energy	EDA	5*1.E-12	Joule per bit
Advertisement energy	Eadv	50*1.E-12	Joule per bit
Range of Communication	RC	20,40,60,80	Meters

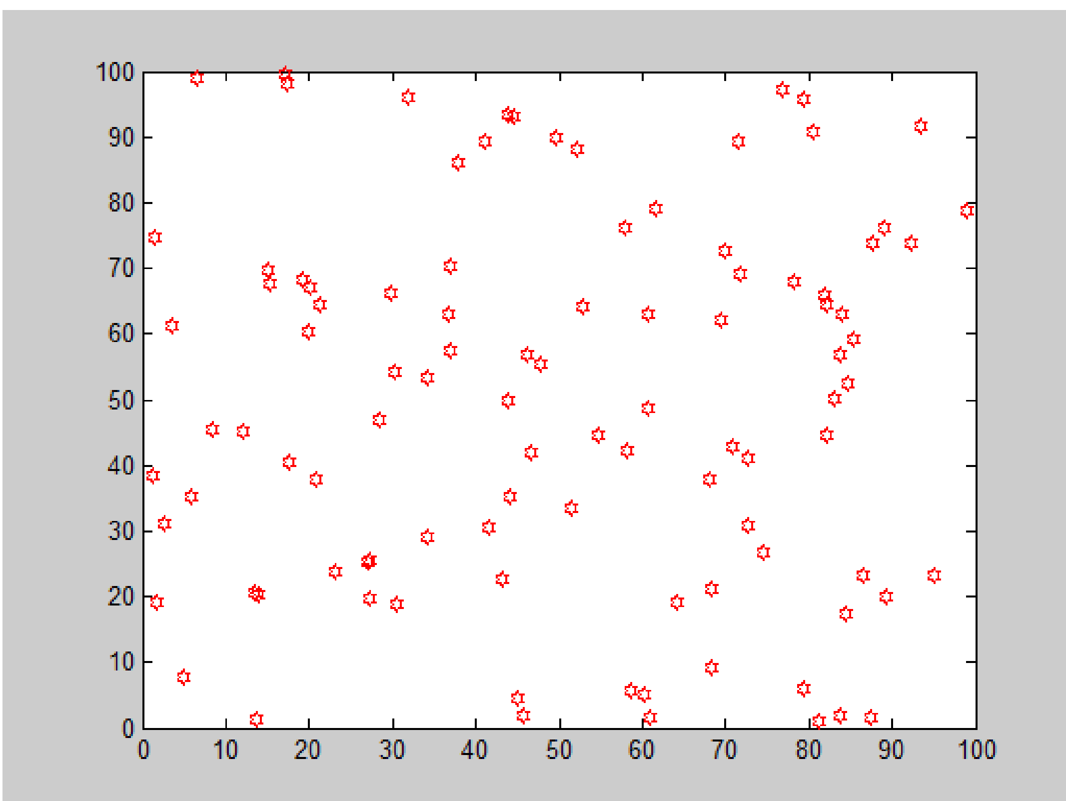


Figure 3.5 - Initial Deployment of 100 Sensors

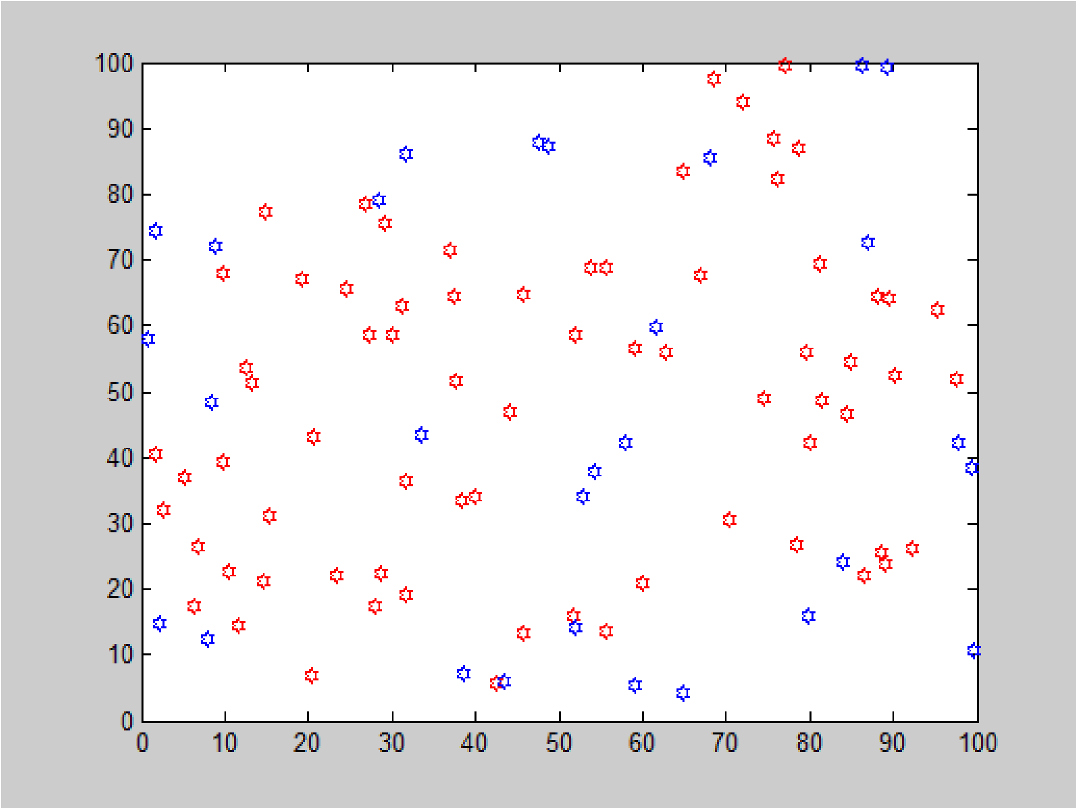


Figure 3.6- Sensor with Cluster Heads

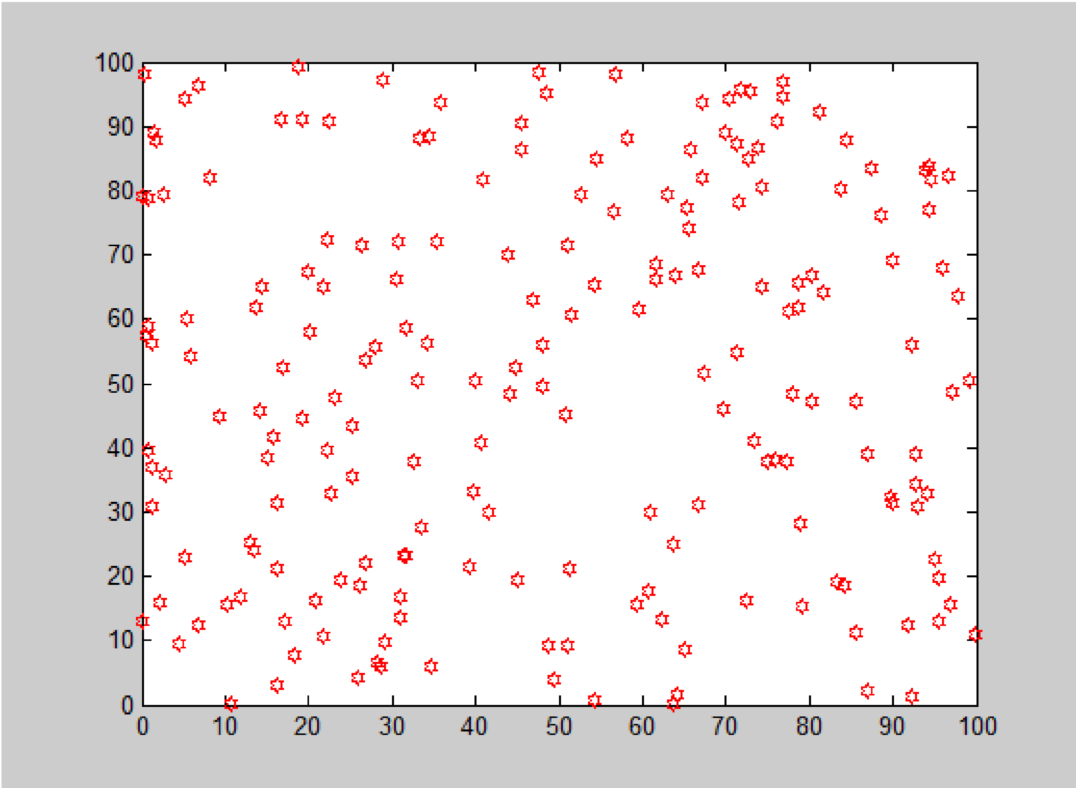


Figure 3.7- Initial Deployment of 200 Sensors

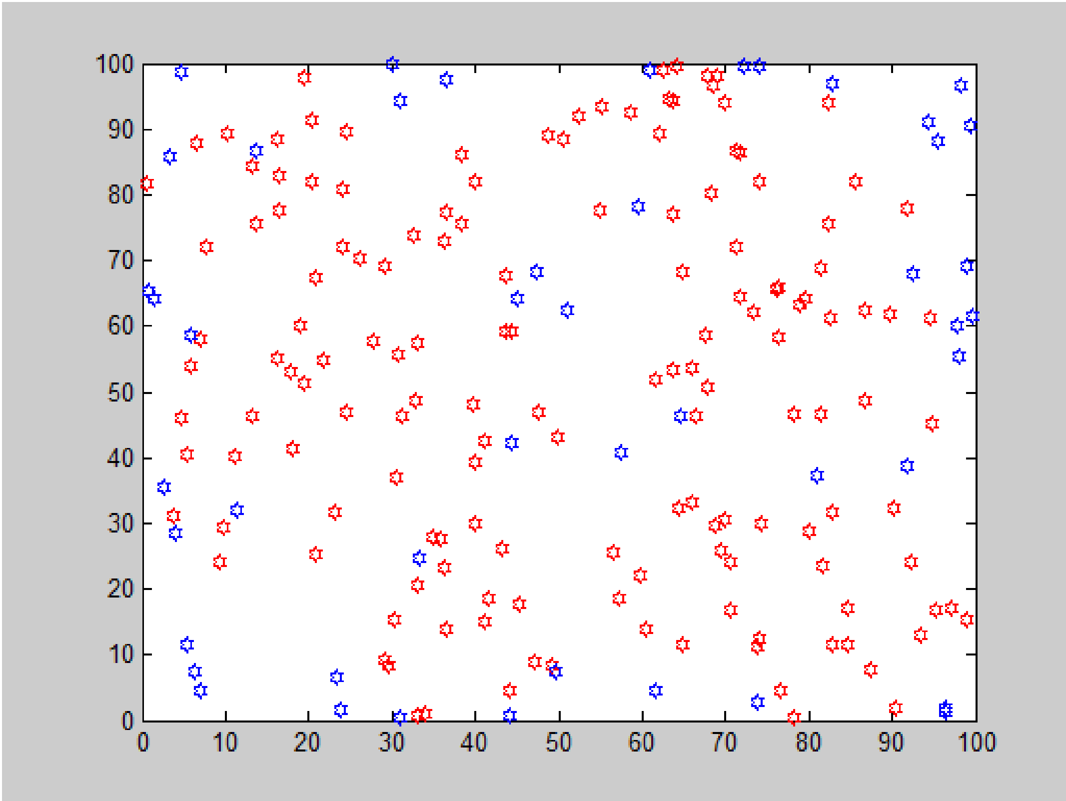


Figure 3.8- Sensor with Cluster Heads

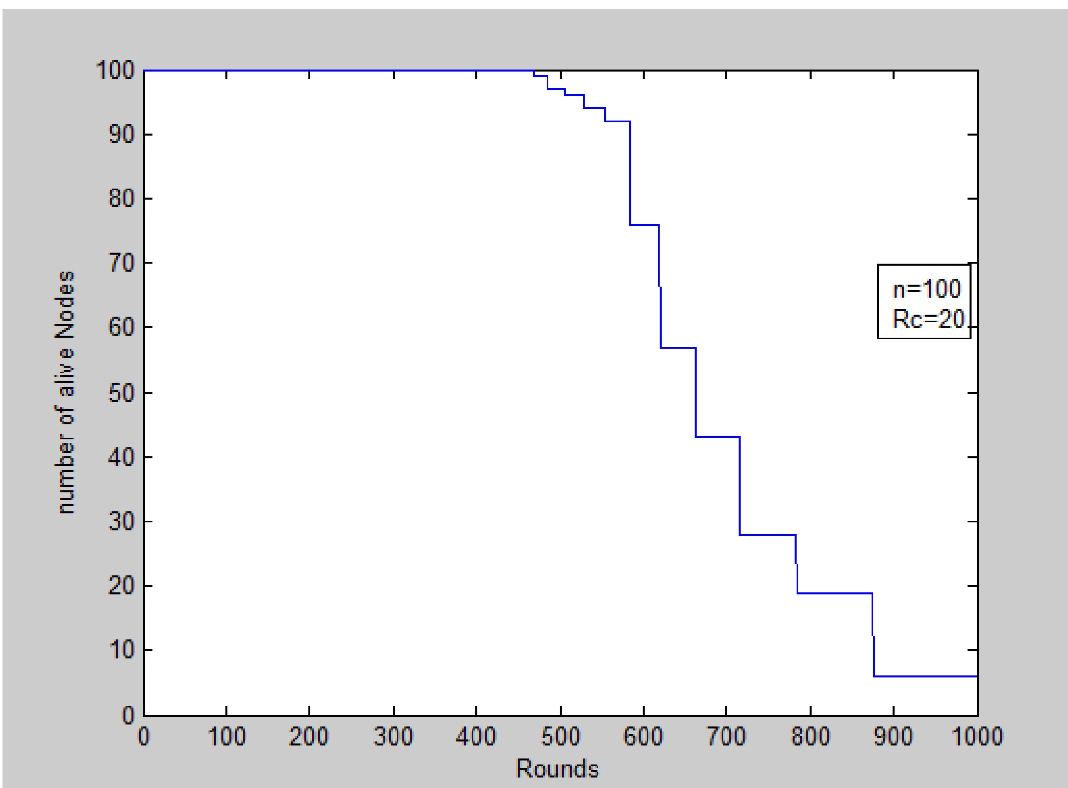


Figure 3.9- Performance of MEDC when no of sensors are 100 and Rc is 20

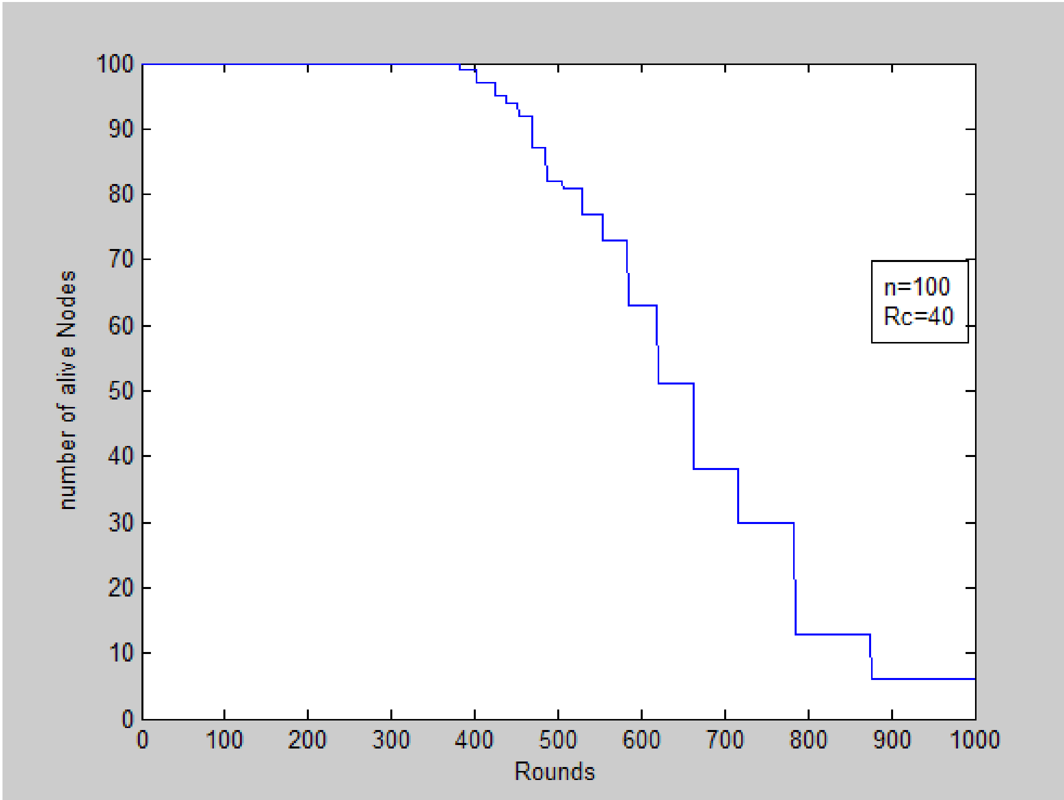


Figure 3.10- Performance of MEDC when no of sensors are 100 Rc is 40.

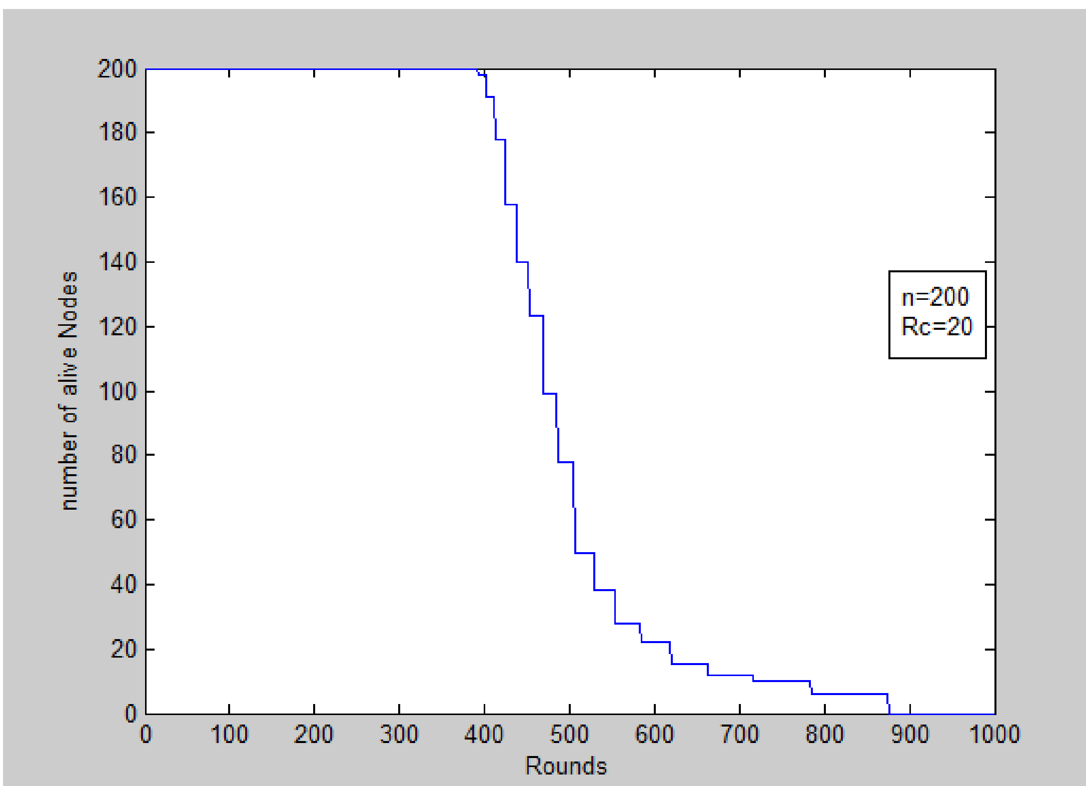


Figure 3.11- Performance of MEDC when no of sensors are 200 and Rc is 20.

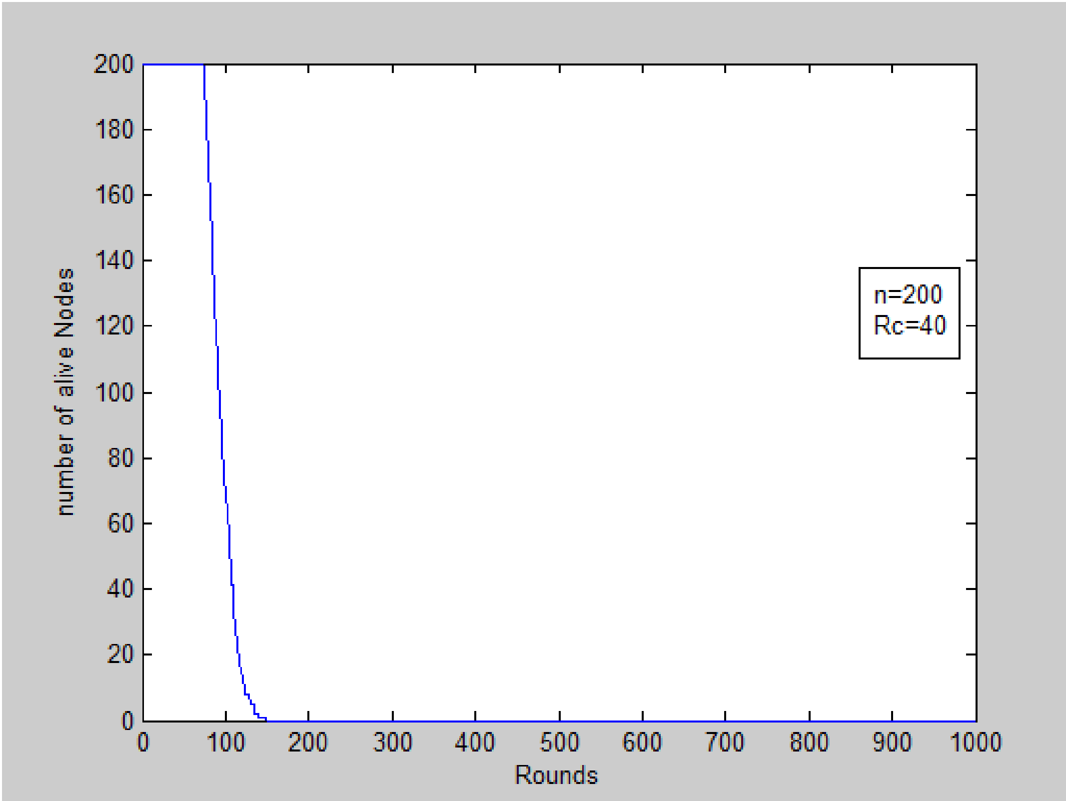


Figure 3.12- Performance of MEDC when no of sensors are 200and Rc=40

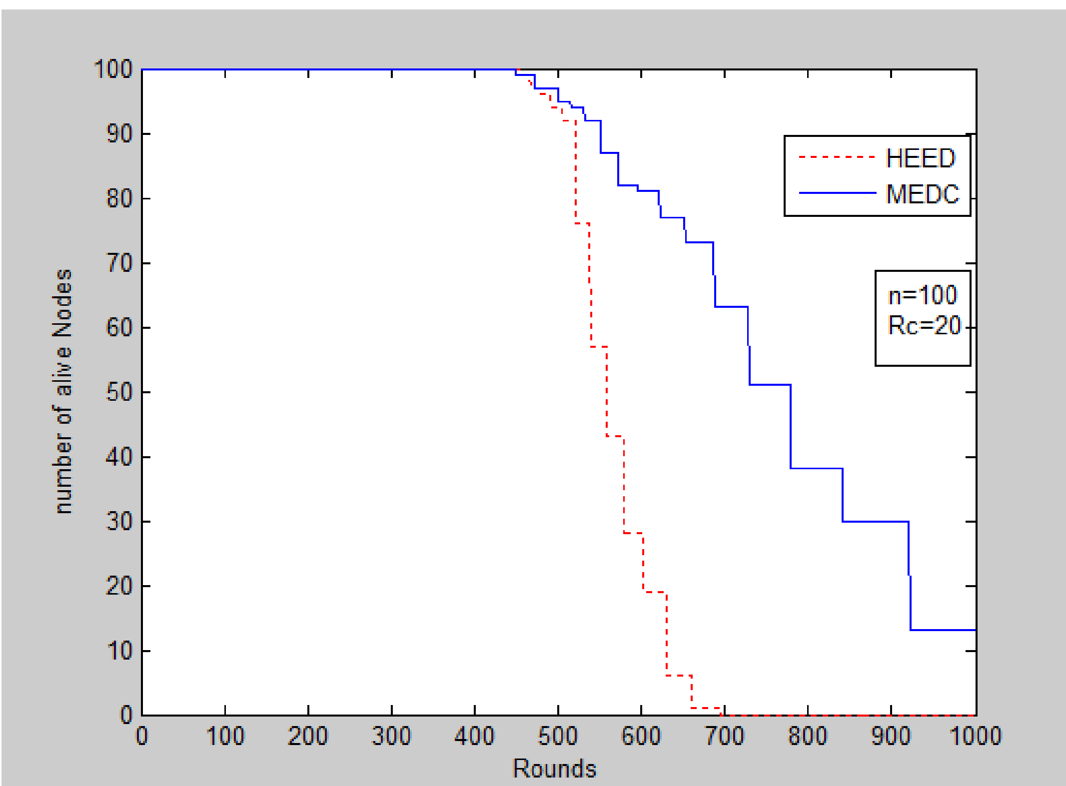


Figure 3.13- Comparative Results of MEDC and HEED number of sensors are 100 Rc is 20

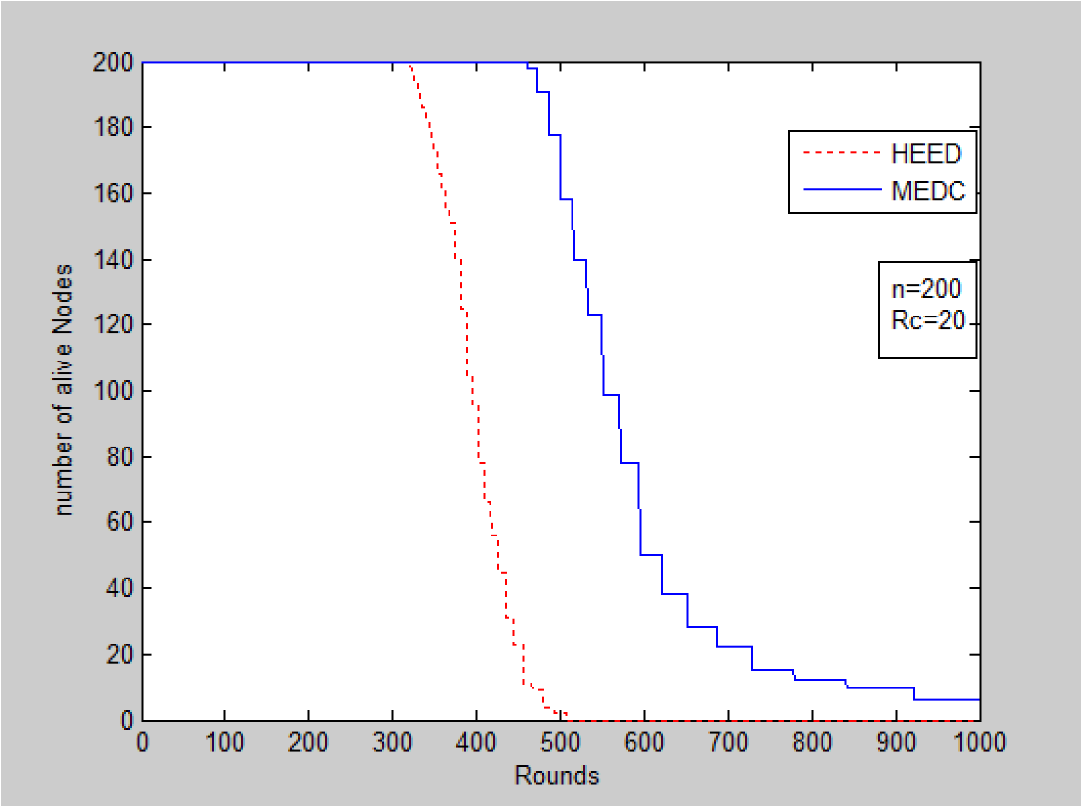


Figure 3.14- Comparative Results of MEDC and HEED number of sensors are 200 Rc is 20.

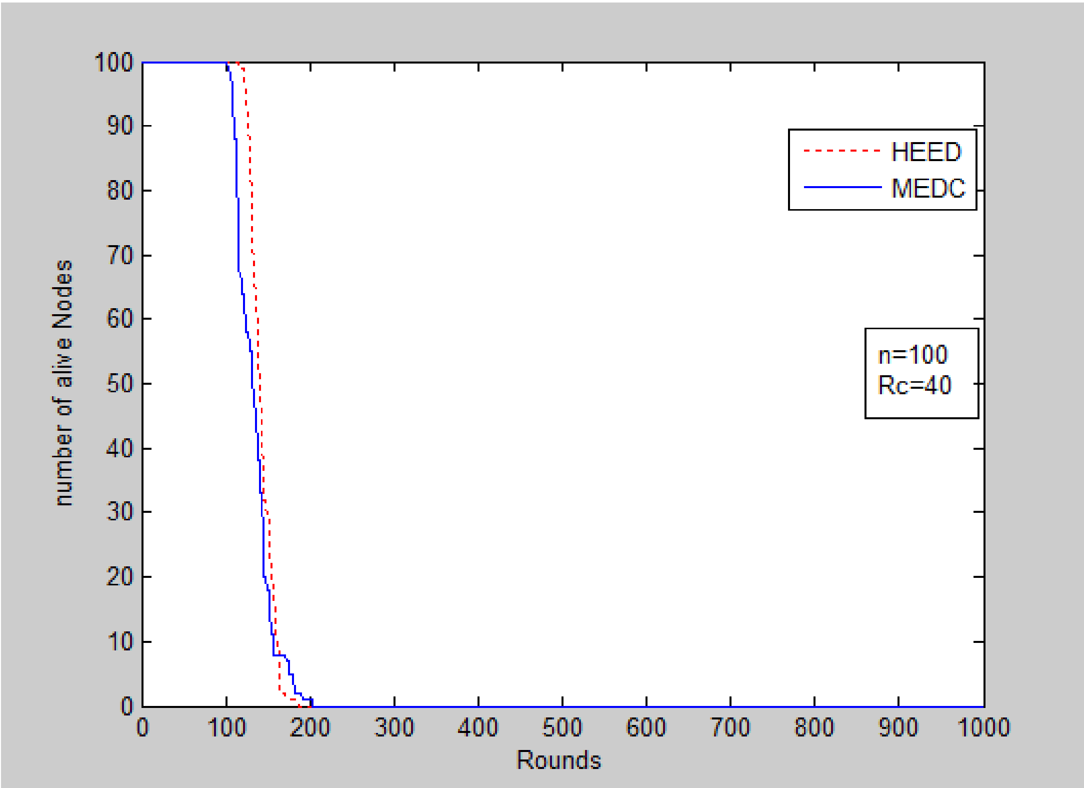


Figure 3.15- Comparative Results of MEDC and HEED number of sensors are 100 Rc is 40

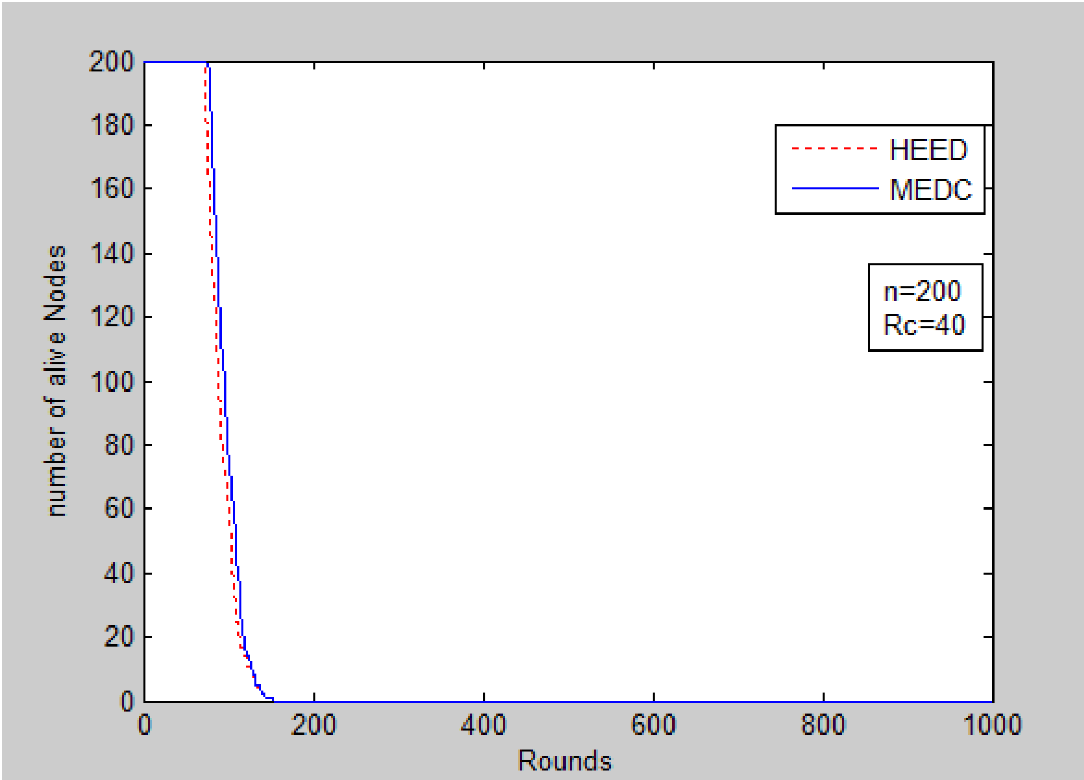


Figure 3.16- Comparative Results of MEDC and HEED number of sensors are 200 Rc is 40

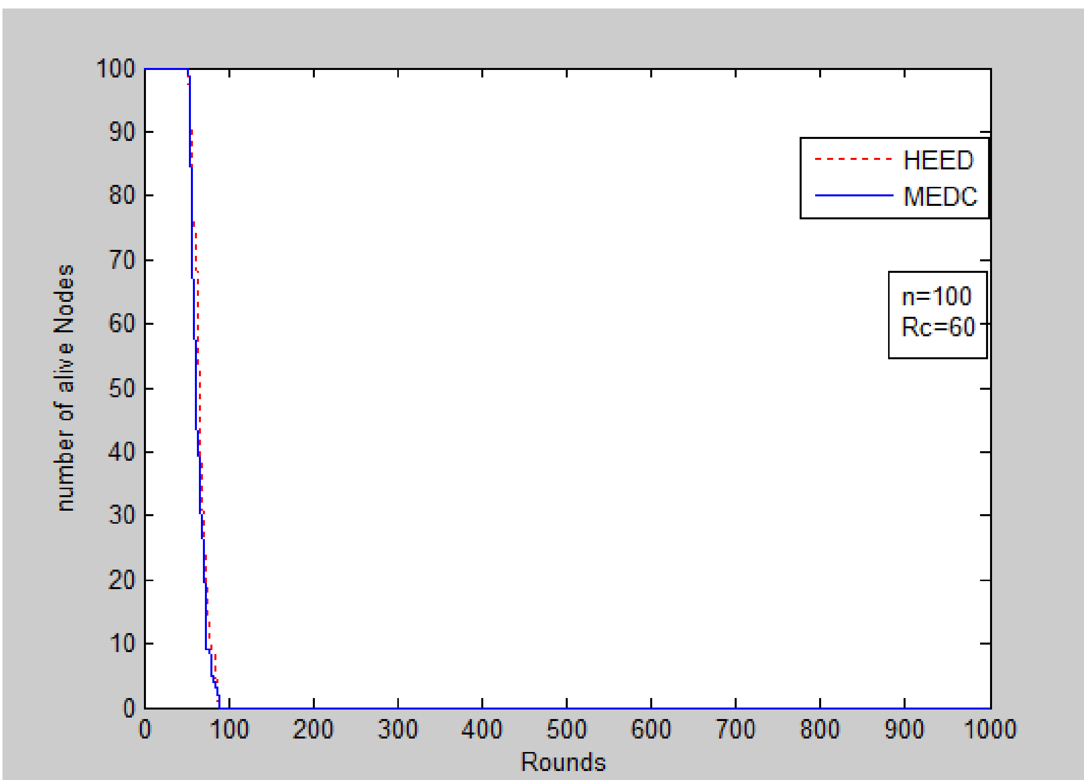


Figure 3.17- Comparative Results of MEDC and HEED number of sensors are 100 Rc is 60

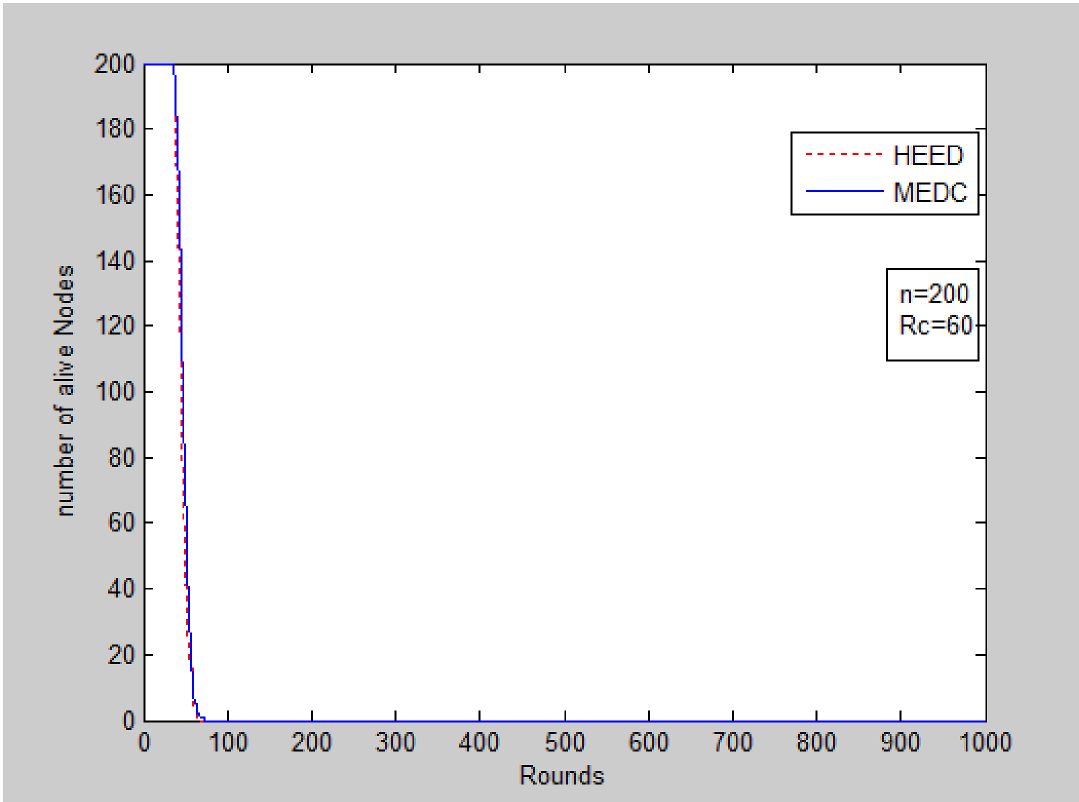


Figure 3.18- Comparative Results of MEDC and HEED number of sensors are 200 Rc is 60

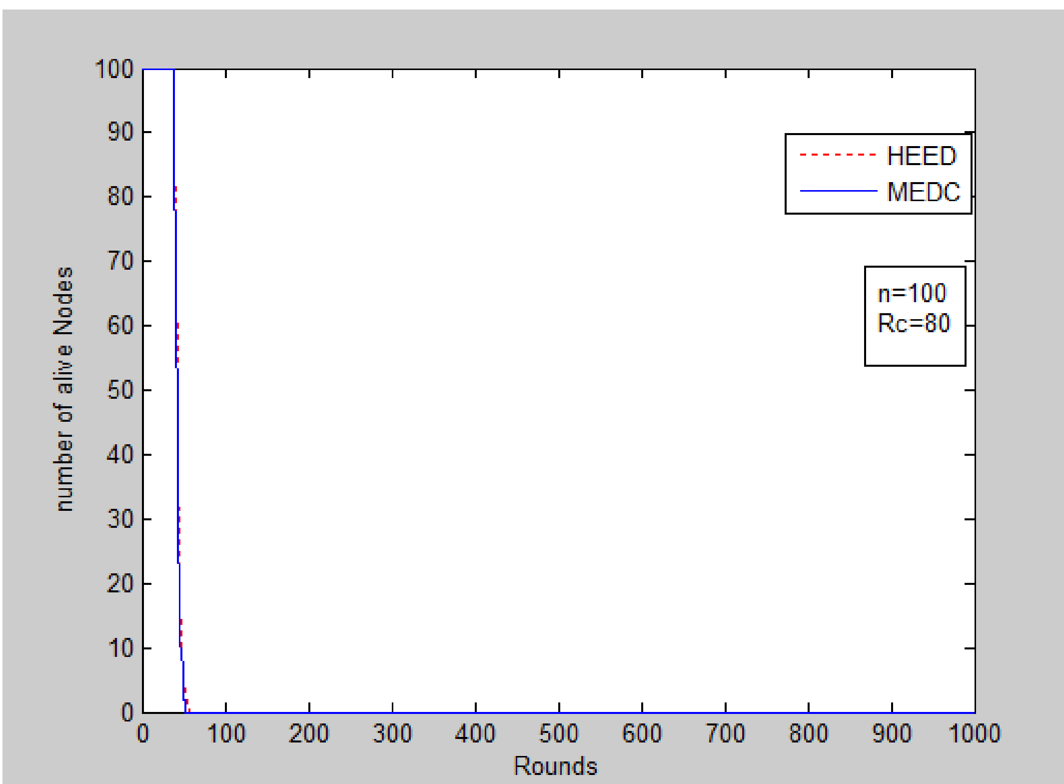


Figure 3.19-Comparative Results of MEDC and HEED number of sensors are 100 Rc is 80

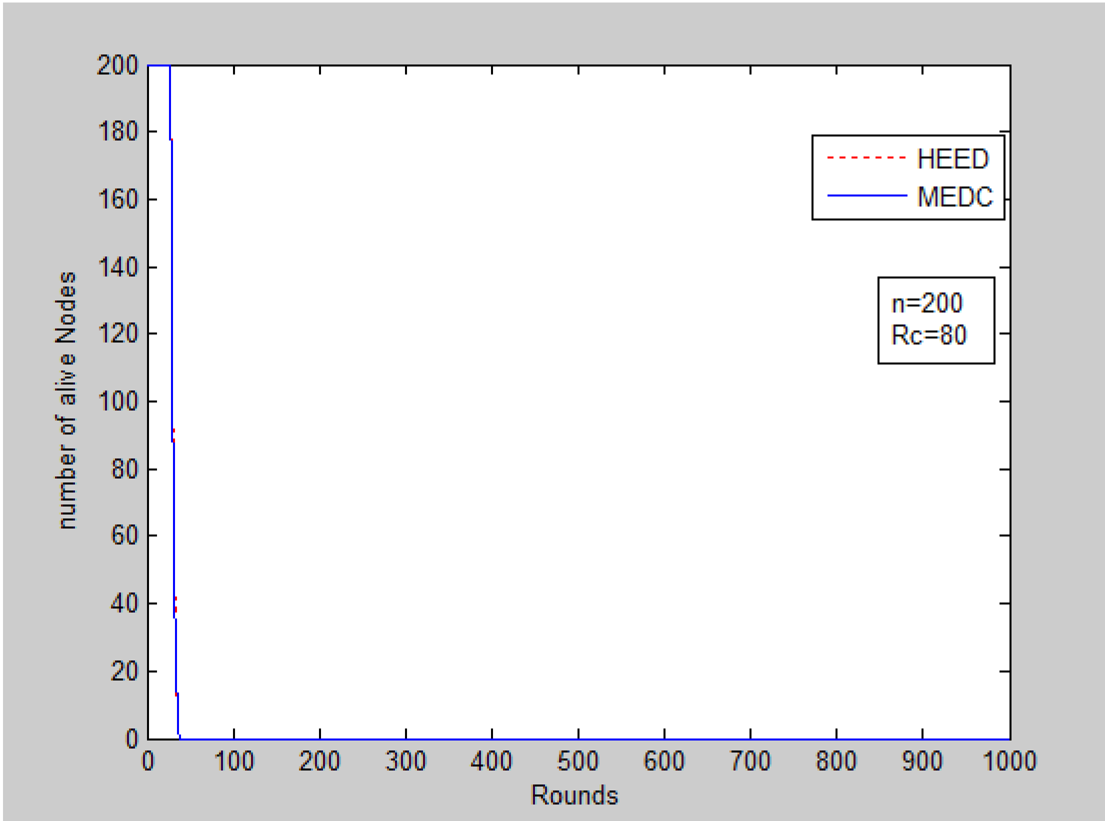


Figure 3.20- Comparative Results of MEDC and HEED number of sensors are 200 Rc is 80

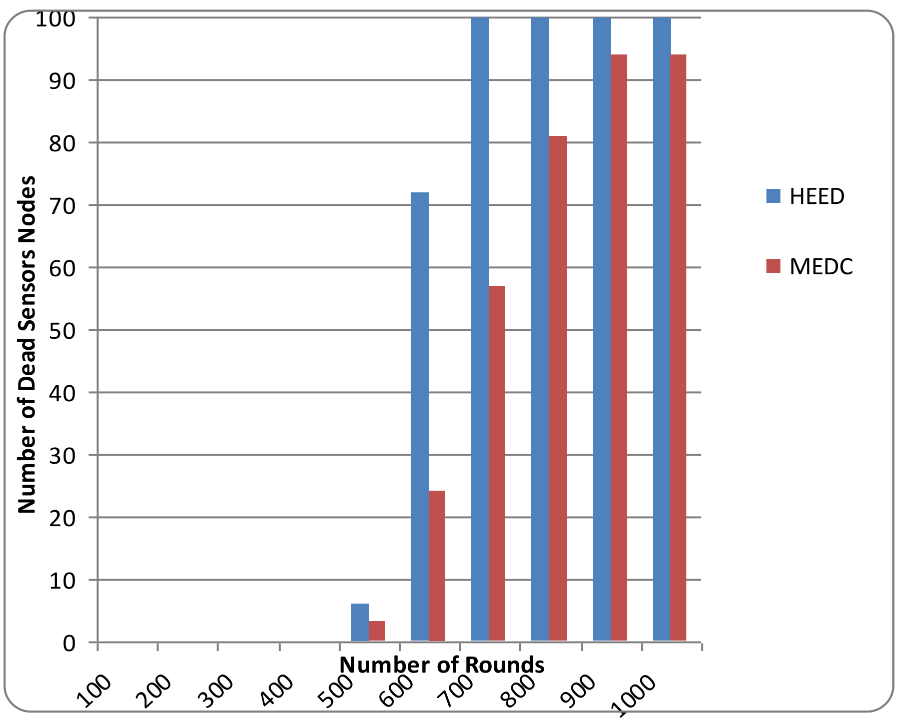
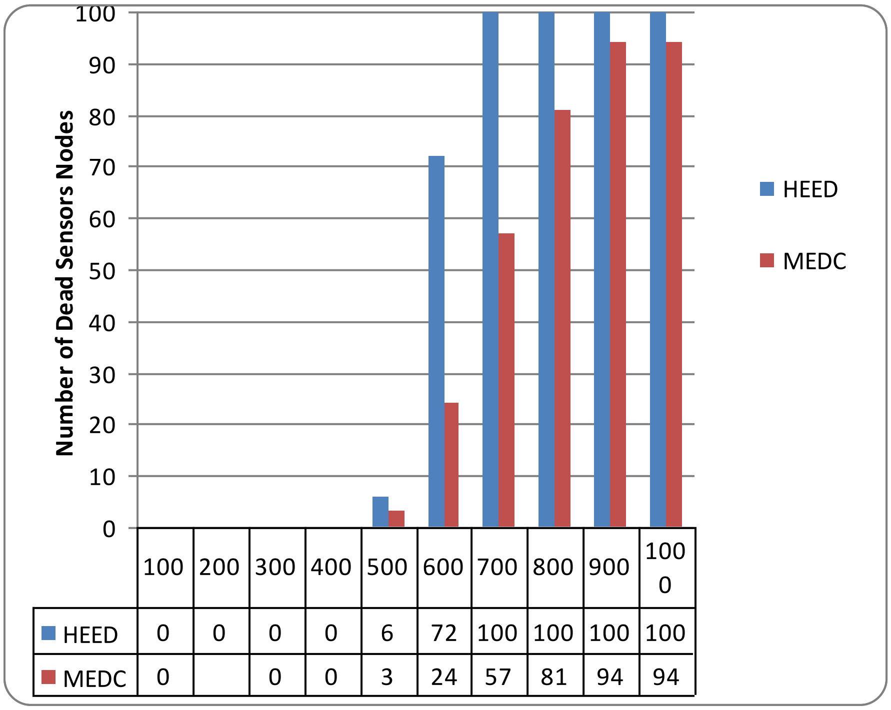


Figure 3.21-Comparison of No of dead sensors V/S Rounds when n=100 and Rc=20



	100	200	300	400	500	600	700	800	900	1000
HEED	0	0	0	0	6	72	100	100	100	100
MEDC	0		0	0	3	24	57	81	94	94

Figure 3.22-Comparison of No of dead sensors V/S Rounds when n=100 and Rc =20 along with detailed table.

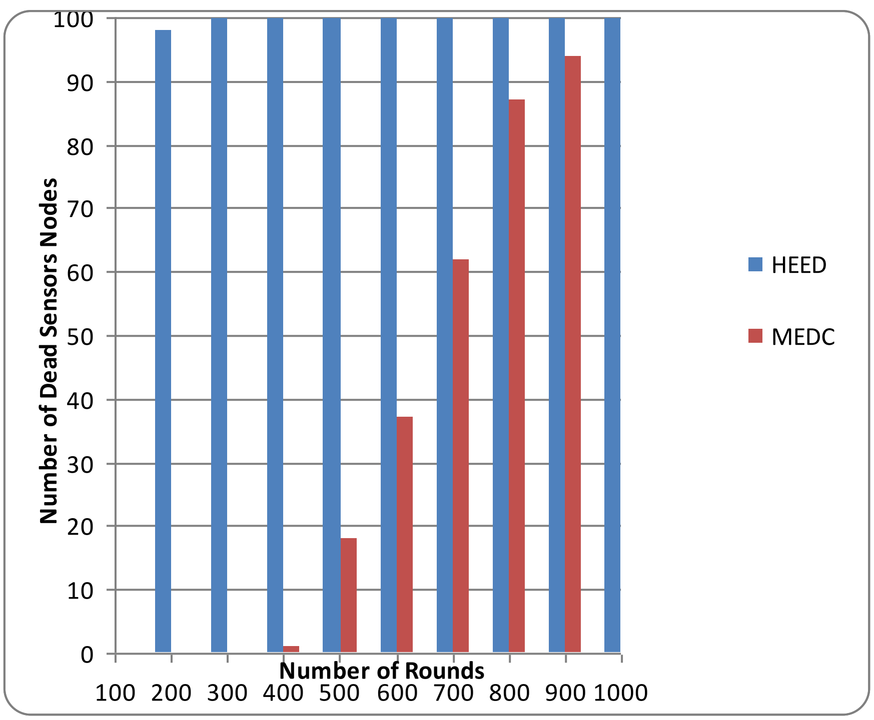


Figure 3.23-Comparison of No of dead sensors V/S Rounds when n=100 and Rc=40

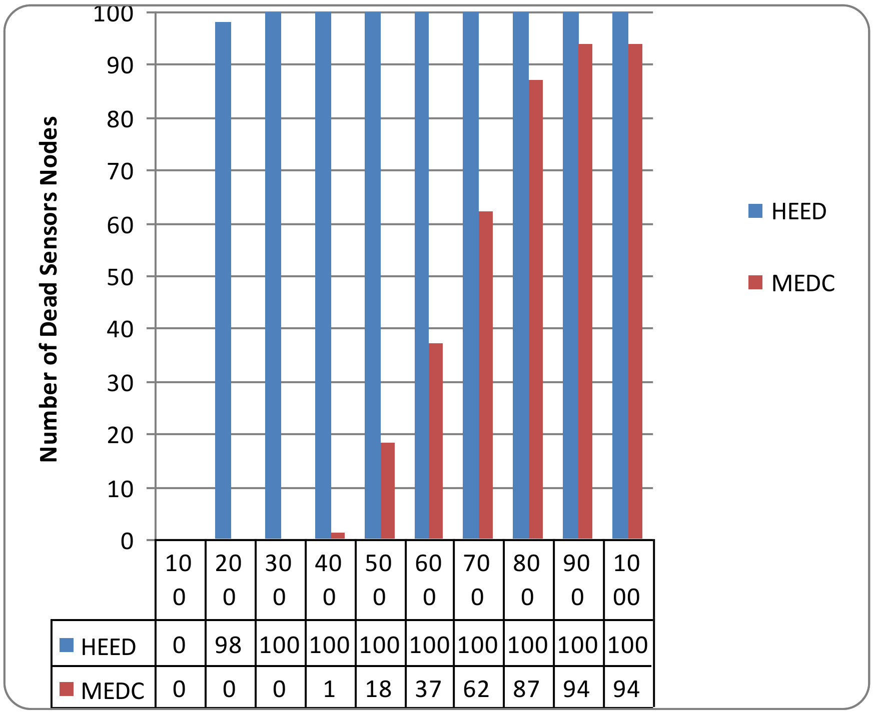


Figure 3.24-Comparison of No of dead sensors V/S Rounds when n=100 and Rc =40 along with detailed table

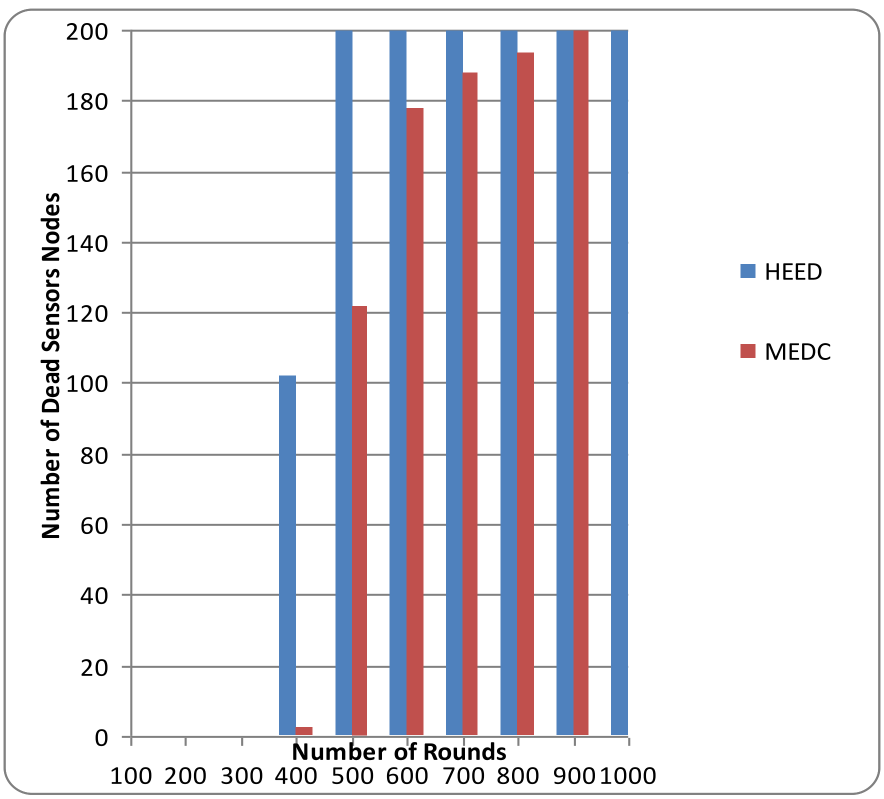
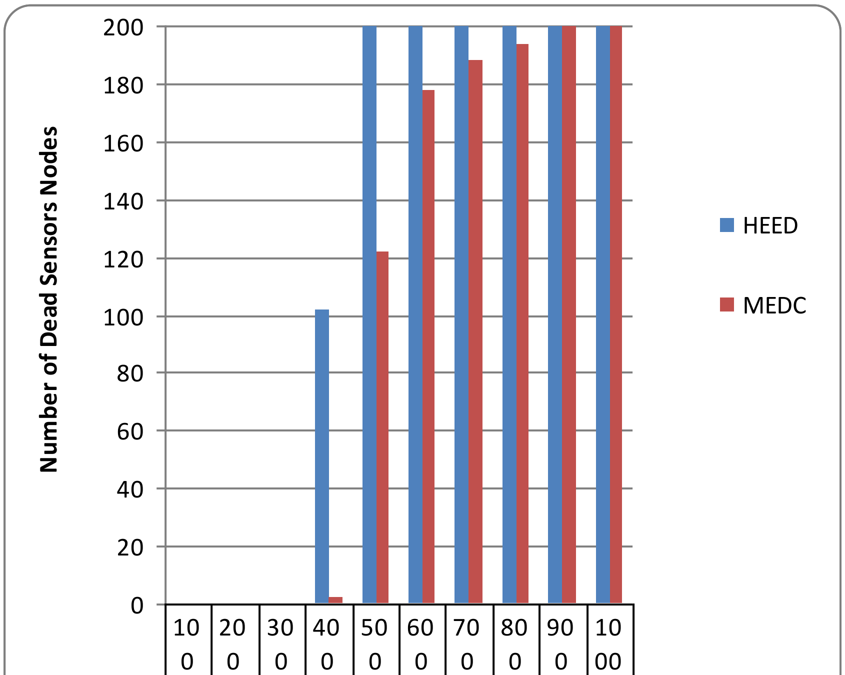


Figure 3.25-Comparison of No of dead sensors V/S Rounds when n=200 and Rc=20



	100	200	300	400	500	600	700	800	900	1000
HEED	0	0	0	102	200	200	200	200	200	200
MEDC	0	0	0	2	122	178	188	194	200	200

Figure 3.26-Comparison of No of dead sensors V/S Rounds when n=200 and Rc =20 along with detailed table

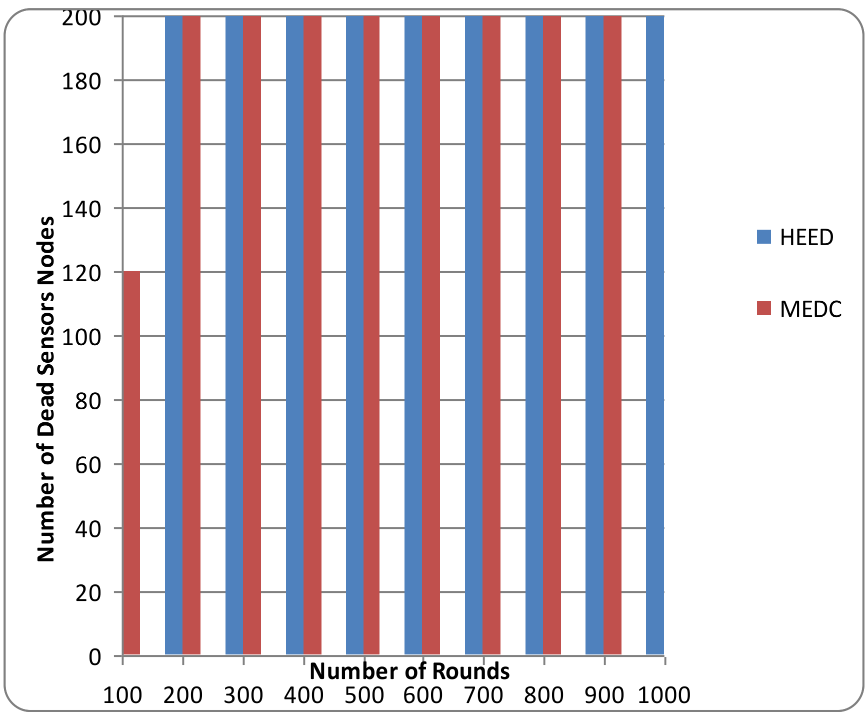
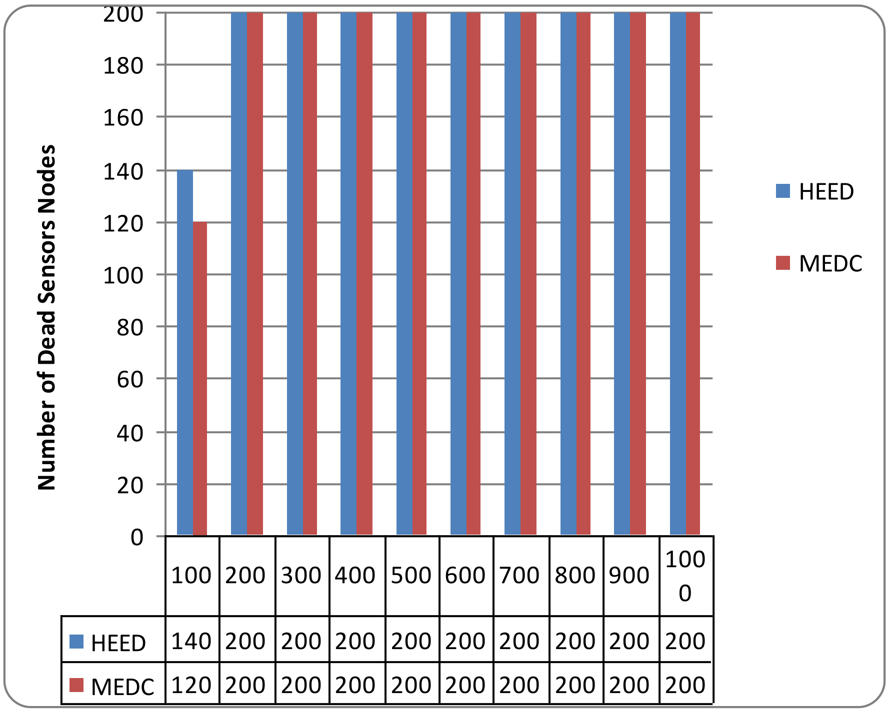


Figure 3.27-Comparison of No of dead sensors V/S Rounds when n=200 and Rc=40



	100	200	300	400	500	600	700	800	900	1000
HEED	140	200	200	200	200	200	200	200	200	200
MEDC	120	200	200	200	200	200	200	200	200	200

Figure 3.28-Comparison of No of dead sensors V/S Rounds when n=200 and Rc =40 along with detailed table

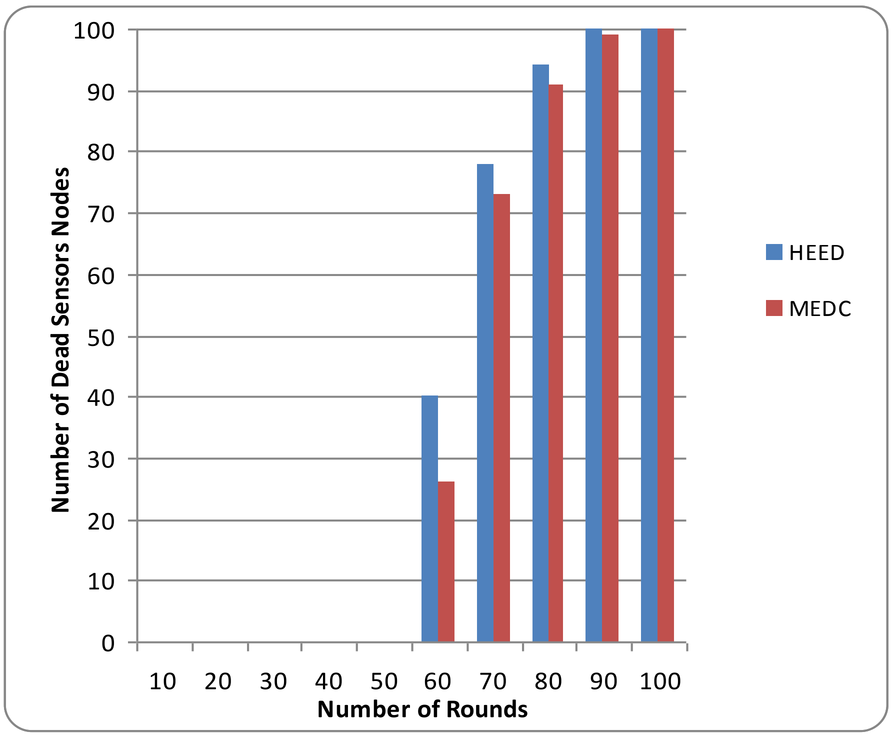


Figure 3.29-Comparison of No of dead sensors V/S Rounds when n=100 and Rc=60

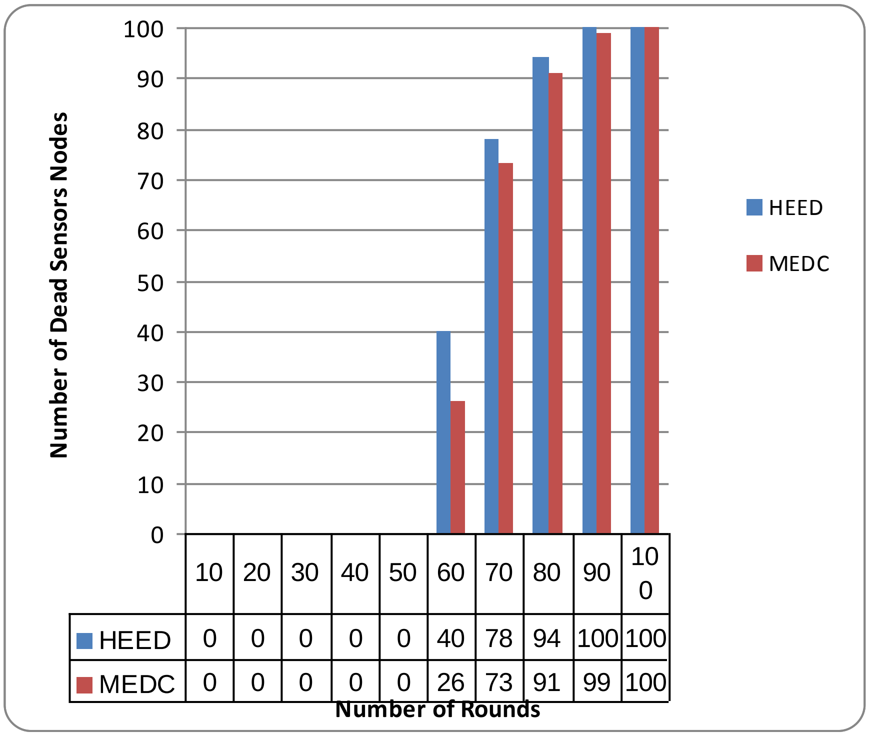


Figure 3.30-Comparison of No of dead sensors V/S Rounds when n=100 and Rc =60 along with detailed table

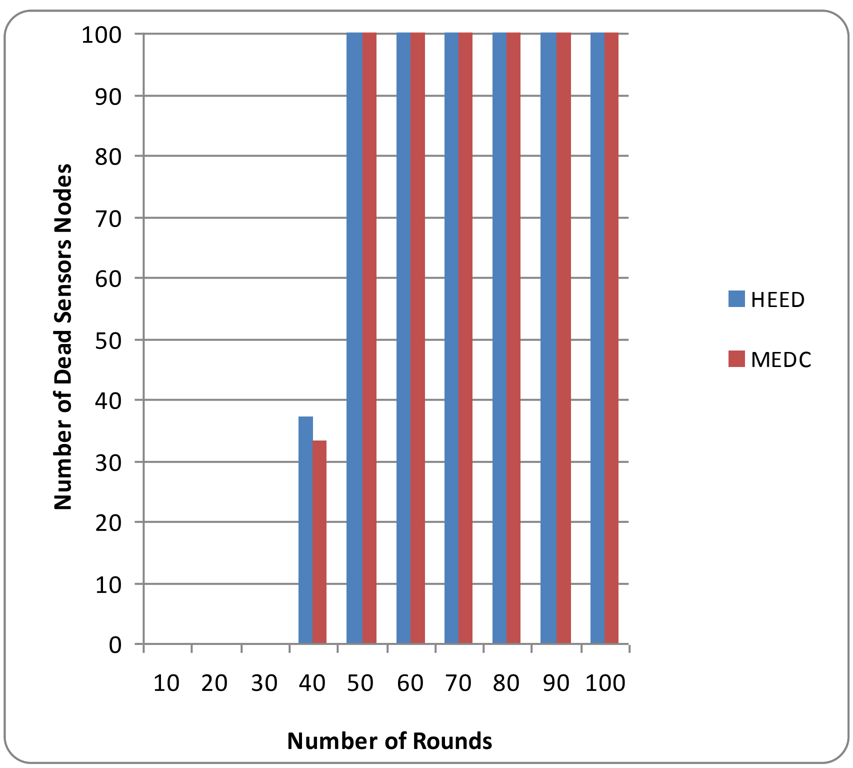
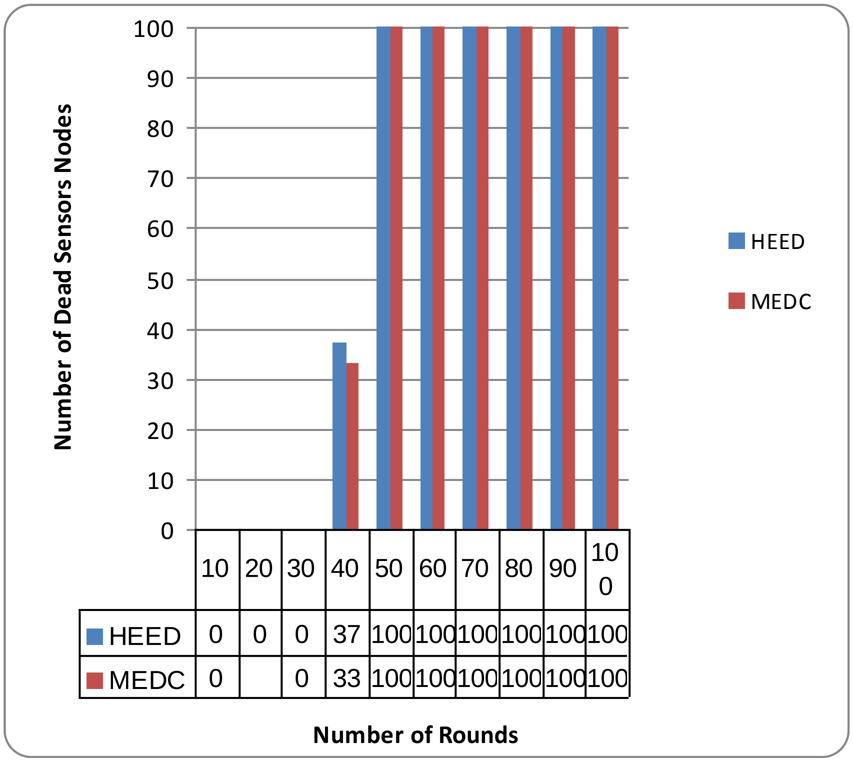


Figure 3.31-Comparison of No of dead sensors V/S Rounds when n=100 and Rc=80



	10	20	30	40	50	60	70	80	90	100
HEED	0	0	0	37	100	100	100	100	100	100
MEDC	0		0	33	100	100	100	100	100	100

Figure 3.32-Comparison of No of dead sensors V/S Rounds when n=100 and Rc =80 along with detailed table

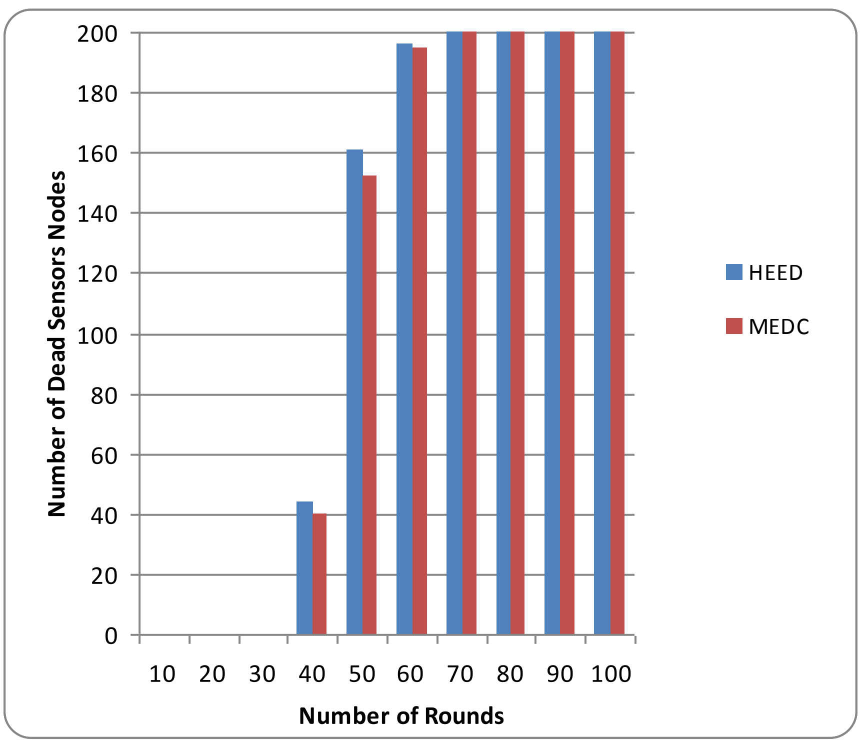
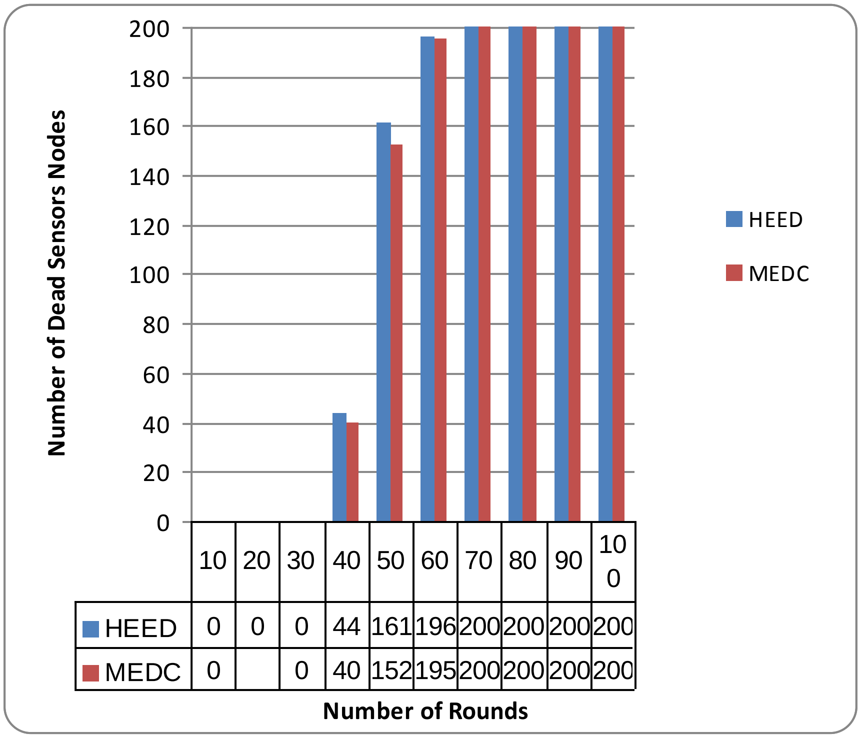


Figure 3.33-Comparison of No of dead sensors V/S Rounds when n=200 and Rc=60



	10	20	30	40	50	60	70	80	90	100
HEED	0	0	0	44	161	196	200	200	200	200
MEDC	0		0	40	152	195	200	200	200	200

Figure 3.34-Comparison of No of dead sensors V/S Rounds when n=200 and Rc =60 along with detailed table

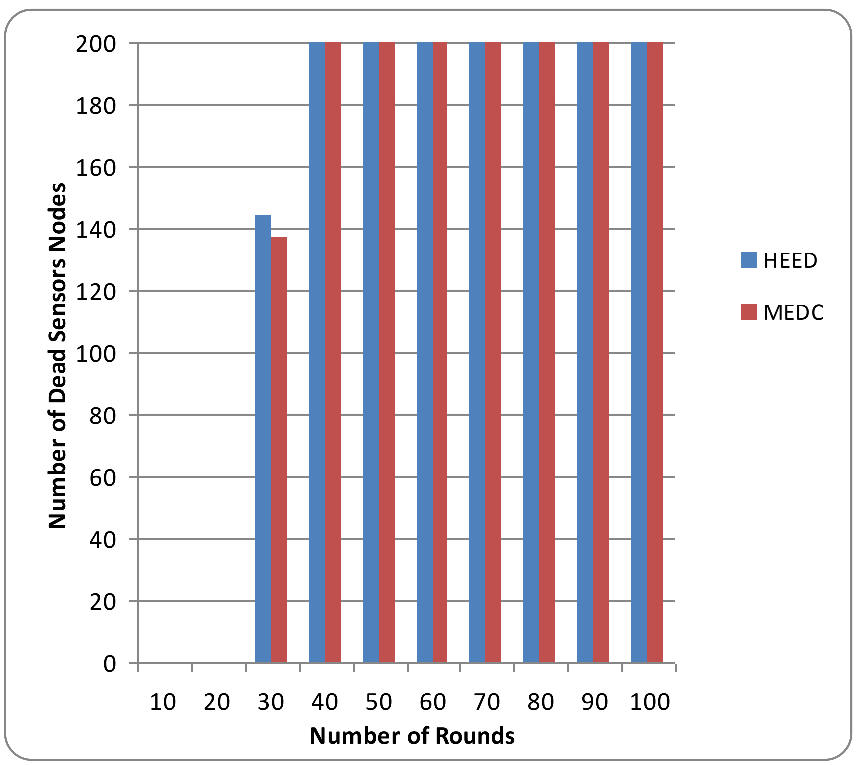
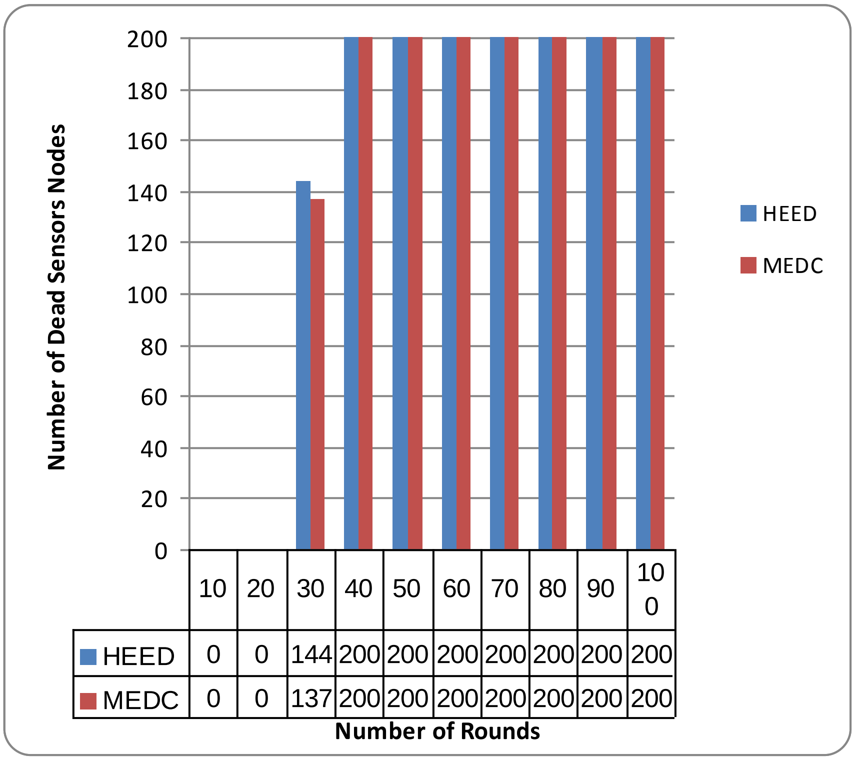


Figure 3.35-Comparison of No of dead sensors V/S Rounds when n=200 and Rc=80



	10	20	30	40	50	60	70	80	90	100
HEED	0	0	144	200	200	200	200	200	200	200
MEDC	0	0	137	200	200	200	200	200	200	200

Figure 3.36-Comparison of No of dead sensors V/S Rounds when n=200 and Rc =80 along with detailed table

3.7 Results Discussions

MATLAB results of MEDC are given from figures 3.5 to figure 3.20. Figure 3.5 is shown result for initial deployment of one hundred sensors over a field. Random function is used for deployment. Figure 3.6 is result of MATLAB shown here for scenario when some sensors under their range become cluster heads. Blue nodes are cluster heads and red nodes are sensors which will act as cluster members under clusters heads from their range of communication. Figure 3.7 and figure 3.8 are results same as like 3.5 and 3.6 respectively but changed n (number of sensor nodes are 200). Figure 3.9 is shown for performance of MEDC on parameter n=100 and range of communication is taken as 20. Performance is measured in terms network life time. Network life time is measured as number of alive nodes per round. In result graph Y axis represents alive nodes and X axis represents rounds. Figures 3.10, 3.11, 3.12 are performance analysis of MEDC by varying values of n and Rc (Range of communication) n is varied on two values 100 and 200. Rc is also varied by two values one is 20 and another is 40. Figures 3.13-3.20 are comparative results of MEDC with HEED. Comparative results are also taken on n value 100 and 200. Rc value is varied from 20, 40, 60 and lastly 80. It is concluded from results that MEDC works better in contrast of HEED. On parameters n=100 and Rc=20 it gives 33.3 improvement in network life time. As parameter are changed, graph of MEDC come close in performance to HEED.

Reason for this decrement in graph is increase cost of communication. When we increase n from 100 to 200 or Rc from 20 to 40, 60 or 80 then their will be more no of sensors under range of communication to each other. More sensors under range more advertisements which mean more cost of communication which result decrease in performance. Figure from 3.21 to 3.36 show results of dead nodes corresponding to rounds. X axis represent rounds and y axis represent number of dead sensor nodes. Up to Graph 3.28 we calculate dead nodes are taken after every hundred rounds and this process is repeated on varying both parameters one is number of sensors deployed and second parameter is range of communication. From figure 3.29 onwards we have calculated dead nodes after every ten rounds as from performance analysis of network life time we came to know that all nodes become dead before 100 rounds for all three protocols that's why for near check we calculated after ten rounds and plotted graph up to hundred rounds on x axis

3.8 Conclusion

In wireless sensor networks battery life saving is important aspect. Clustering protocols provide a solution to it. This work has proposed a new clustering protocol named MEDC. MEDC is distributive protocol works on message passing system. It select cluster on mutual exclusion algorithm basis. MEDC protocol is simple in nature and select cluster heads to those sensors which have maximum residual energy within range of communication.

On similar network assumption MEDC perform better than HEED because it consider directly residual energy up to last iteration for next iteration's cluster heads selection in contrast of HEED where residual energy is considered for probability calculation. It is also found in worst cases when the sensors with long range of communication are deployed poorly large in number over small region; which may increase in between message communication among the sensors even in these cases MEDC perform equal to HEED.

EFFICIENT DISTRIBUTED CLUSTERING

4.1 Introduction

Clustering protocols are important because they have major role in improving network life to make networks more sustainable, scalable and efficient [64]. From survey it is found that, what are different working parameters and features of various clustering protocols? In previous chapter we proposed protocol named as Mutual Exclusive Distributive Clustering (MEDC). Section 4.2 discuss about background of MEDC and HEED protocol. HEED protocol work on Ch_{prob} and Snbr on the second side MEDC protocol work on mutual exclusion distributed algorithm. Working parameter of MEDC was residual energy and range of communication. Experimental results showed MEDC perform better than HEED in most of times, even in challenging condition it perform equal but not less than HEED. One part of background is about MEDC, algorithm can be referred from previous chapter.

Section 4.3.1 will discuss some network model assumptions related to our work 4.3.2 will discuss Radio model of communication for protocol. Section 4.3.3 will present gist of MEHEED protocol. This work is carried to add benefits of both two protocols one is MEDC and other is HEED. MEHEED protocol will take three parameters for its working. Out of them one parameter will be Ch_{prob} as like HEED does. Next two parameters will be residual energy and range of communication as like MEDC does. MEHEED will firstly try to select cluster heads on Chprob basis.

Section 4.4 will be about algorithm which is diagrammatically represented as flow chart in section 4.5. Section 4.6 is about experimental evaluation. We have simulated MEHEED protocol on MATLAB. Results are taken on varying two parameters among them one is number of sensor nodes and second is range of communication. For simulation we had taken two values for sensors one is 100 another is 200. For Range of communication we had simulated on four values 20, 40, 60 and 80. Last section 4.7 is about result discussion and section 4.8 will discuss conclusion of this work. Conclusion part also discuss advantage and drawbacks of this proposed work.

4.2 Background

To develop MEHEED we have tried to merge two protocols one is HEED and second is MEDC [80]. Both protocols are distributive in nature. Here we want to present review of these two protocols. HEED is distributive in nature and power base protocol [10]. HEED protocol periodically selects cluster heads according to a hybrid of the node residual energy and intra-cluster communication cost. HEED has an advantage that it does not require special node capabilities, such as location-awareness, or special node distribution etc. Cluster head selection in HEED protocol is primarily based on the remaining energy of each sensor node. Remaining energy is estimated current energy in the node. Secondary clustering parameter that authors has considered is intra-cluster communication cost.

The algorithm is divided into three phases: first one is initialization phase, second one is repletion phase followed by third phase that is finalizing phase. In first phase broadcast cost is calculated according to nodes under range of communication and also cluster head probability for each sensor node is calculated. Second phase is repeat phase; this phase decides tentative cluster heads. If any node is having probability equal to one then this sensor node will be declared as final cluster heads otherwise least cost head is taken as tentative cluster head. Cluster heads are finalized in last phase. And sensor nodes become cluster member under least cost reachable cluster head. HEED protocol also ensure that it will terminates in a constant number of iterations, independent of network diameter. Here is algorithm of HEED protocol

I. Initialize

1. Snbr ← {v: v lies within my cluster range}
2. Compute and broadcast cost to Ɛ Snbr
3. CHprob ← max (Cprob × Eresidual/ Emax , pmin)
4. is_final_CH ← FALSE

II. Repeat

1. If ((SCH ←{v: v is a cluster head}) != Φ
2. my_cluster head ←least cost(SCH)

```
If (my_cluster _head = NodeID)
            If (CHprob = 1)
                        Cluster_head_msg(NodeID,final CH,cost)
                        is_final _CH ← TRUE
            Else
                        Cluster_head_msg(NodeID, tentative_CH,cost)
ElseIf (CHprob = 1)
            Cluster_head_msg(NodeID,final CH,cost)
            is_final_CH ←TRUE
ElseIf Random (0, 1) <=CHprob
            Cluster_head_msg(NodeID, tentative_CH,cost)
            CHprevious← CHprob
            CHprob ← min(CHprob × 2, 1)
Until CHprevious = 1
```

III. Finalize

```
If (is_final_CH = FALSE)
    If ((SCH ←{v: v is a cluster head}) != Φ
            my_cluster_head← least_cost(SCH)
            join cluster(cluster_head_ID, NodeID)
    Else Cluster_head_msg( NodeID, final CH, cost)
Else Cluster_head_msg(NodeID, final CH, cost)
```

The second protocol we had considered is MEDC which is our first proposed work. Mutual Exclusive Distributive Clustering (MEDC) will form clusters of sensor on basis of mutual exclusion algorithm. In MEDC Clustering protocol the Cluster heads will be chosen in a mutual exclusive way over a range of communication. Under a range of communication (says R_c), only a sensor which is having maximum of residue energy will be the cluster head. The proposed protocol will run in iterations, each iteration follow four steps and the cluster head will be reechoes in succeeding iteration. Under R_c, the Cluster head sensor node will be decided on factor of remaining energy. Sensor node which one is having highest remaining energy among the sensors node under R_c will be chosen as cluster head. The idea of Mutual exclusion works by message passing system we will say it by advertisements. Why we are considering the remaining energy factor , is because sensor's remaining energy will be different for each sensor nodes infect it depend on how much each sensor node spends energy for sending and receiving advertisements up to last iteration. Which in turn how much advertisement depends on how many sensors were under range of communication (R_c). Algorithm can be referred from section3.4.

4.3 MEHEED

4.3.1 Network Model assumption

- Sensor nodes are of homogeneous type means they have the same capabilities and resources like battery power etc.
- Sensor nodes are stationary deployed in local region for monitoring and the data sink is located far from the sensing field.
- Sensor nodes are considered as location–unaware that means they are not equipped with GPS or other similar equipment.
- Periodically the recently sensed data and information by all nodes are gathered and sent to the data sink after aggregation.
- Each sensor node is assigned a unique identifier (ID). Each sensor will be assigned initial probability Cprob which is used to calculate Chprob.
- The links are assumed to be symmetric. Communication can be bidirectional.

4.3.2 Radio model Equations

Sensors energy is dissipated on transmission and receiving activity along with sensing [61]. If any sensor want to transmit k bits to a node located at distance d. Then energy dissipation is calculated by equation 4.1

- Transmission Energy Dissipation

- Etx (k) =k* Eelec. + Etx_amp (k, d) (4.1)
- Etx_amp (k, d) = k*d^2*Efs (4.2)
- Etx_amp (k, d) = k*d^4*Emp (4.3)

Where Etx represent transmission energy will be calculated from electron energy (Eelec) and amplification energy (Etx_amp). Amplification energy consumption also varies depending on free space communication or multipath communication equations 4.2 and 4.3 are shown. Free space energy and multipath energy consumption is represented as Efs and Emp respectively. If any sensor node is receiving k bit from any sensor node then energy dissipated for k bits is represented as Erx

- Receiving Energy Dissipation
  - Erx (k) = k* Eelec (4.4)

4.3.3 Gist of MEHEED

MEHHED protocol is extension work of our first contributory work MEDC protocol. MEHEED protocol is combination of MEDC and HEED protocols. MEDC protocol was working on the parameters of residue energy $E_{residual}$ and range of communication on the other side HEED protocol considers three factors one of them is Ch_{prob} second is Snbr and third is range of communication. The proposed MEHEED protocol will take first parameter same i.e. Ch_{prob} and second parameter will be $E_{residual}$ instead of S_{nbr}, the third factor is same for all three protocols here i.e. Range of communication.

The idea to change the second parameter is; instead of considering previous calculated Snbr which was dependent on remaining energy of starting level, why not to consider $E_{residual}$ that have been recalculated after each iteration. Benefit of this idea will be that recent updated value i.e. $E_{residual}$ will also reflect energy detrainment of previous cluster heads. So decisions will be more accurate. MEHEED protocol adopts benefit of both protocols. When it take first decision sensor's Chprob will be checked, which is calculated in first phase. If Chprob come out equal to one that it will select that particular sensor as cluster head, rest all computations will be simply skipped. if it is not equal to one in that case part of MEDC algorithm works out and selects cluster head which is having maximum residual energy. MEHEED protocol will works in two phases.

First phase will be of initialization and calculations phase is as like HEED. In first phase first of all sensors under the range of communication are queued. On basis of this queue, Communication cost and $E_{residual}$ will be calculated. After that, Ch_{prob} will be calculated on basis of residual energy and predefined C_{prob} as like of HEED protocol.

Second phase will decide the cluster head and the cluster members under clusters. Second phase will decide which sensor will be cluster head. This decision firstly depends on Ch_{prob} after that this decision will depend on the $E_{residual}$ unlike HEED, in which second factor was S_{nbr}. If and sensor node is having Ch_{prob} is equal to one that it will be directly declared as cluster head and sensors under its queue will be cluster members for this. In this case further computations will be skipped both computation time and energy will be saved. In case if there is no sensor node under a range of communication have met first selection criteria then selection will be done according to MEDC protocol i.e. on basis of remaining energy. MEDC protocol will select cluster head to that sensor that is having maximum residual energy in simple way; no tentative cluster heads will be selected as in HEED again complex computations will be skipped .The results have been shown that this idea has given more affective results to save network life. Reader may wish to know

How CHprob can be equal to 1?

As we said CHprob = max (Cprob × Eresidual/Emax, pmin).

Assume sensors may have their energy in range [0.0, 0.2]. 0.0 joule is the minimum energy a sensor can have and 0.2 joule is the maximum. Here Cprob is initialized as 0.5. Pmin is assumed minimum probability equal to 0.1.

Eresidual is remaining energy of sensor say nodes i. Assume node i is having 0.18. Emax will be calculated by average from network energy means calculating total energy of all sensor nodes and dividing by total no of sensor nodes. The fact is when sum will be calculated at that time some sensors might have energy at zero but they are part of network so their number will be calculated in network but energy in sum will be zero. Say we are having 50 nodes and. Let us say we had calculated sum that come out 4.5 joule. Average energy will be assigned to Emax that will be (4.5/50=0.09).

Substituting the values in equation CHprob = max (0.5 × 0.18/0.09, 0.1) =1

Table 4.1- Working Factors for Clustering

HEED	MEDC	MEHEED
1. Ch_{prob} 2. Snbr (cost) 3. Range of Communication	1. Eresidual 2. Range of Communication	1. Ch_{prob} 2. Eresidual 3. Range of Communication

4.4 Algorithm MEHEED

Two phase protocol

i- Sensor node

Rf- Range of frequency

v- Nodes under Rf

Qi-Queue of sensor i

Cprob- Predefined probability of each sensor

CHprob- Probability of being cluster head

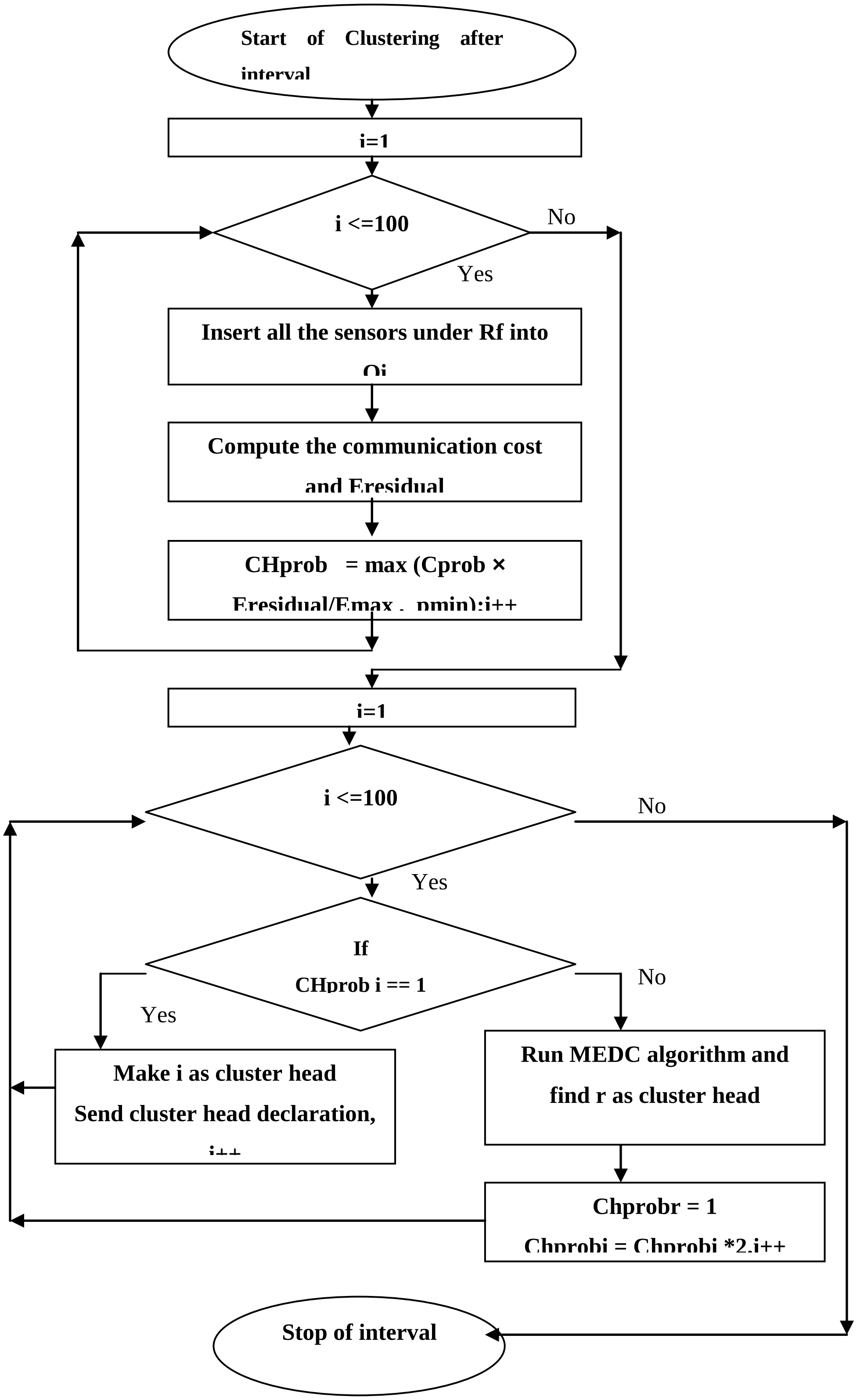
Phase I

```
1.      For each sensor i
2.              Qi <-- v: v under of Rf
3.                      Compute the communication cost of i after investing on Qi. And
                        find the Eresidual
4.              CHprob   = max (Cprob × Eresidual/Emax , pmin)
```

Phase II

```
1. For each sensor i
2.      If (Chprobi == 1)
3.              cluster_head_declaration message
4.      else if
5.                      Advertise Epresent_i each j  within  Qi
6.                      Queue all incoming advertisements from sensors j into adv_Qi
7.      Compare Epresent_i  with  Epresent_j in adv_Qi
8.      If (Epresent_i >= Epresent_j)
9.                      cluster_head_declaration message
10.      else
11.             find  j in adv_Qi having Epresent_j  >= Epresent_i
12.                     compare this Epresent_j  with other sensors in adv_Qi and find the
                        highest energy sensor let  its r
13.     do Chprobr = 1
14.     do Chprobi = Chprobi *2
```

4.5 Flow Chart



4.6 Experimental Evaluation

For simulation of MEHEED on MATLAB we had taken following parameters. Shown in Table 4.2

Table 4.2- Simulation Parameters for MEHEED

Parameters	**Abbreviation**	**Values**
Random field x axis	Xm	100 meter
Random field y axis	Ym	100 meter
Initial energy of sensor	eo	0.05 Joule
Total number of sensor	n	100,200
Transmission energy per bit	Etx	50*1.E-9 Joule
Receiving energy per bit	Erx	50*1.E-9 Joule
Free space energy per bit	Efs	10*1.E-12 Joule
Data aggregation energy per bit	EDA	5*1.E-12 Joule
Advertisement energy per bit	Eadv	50*1.E-12 Joule
Range of Communication	Rc	20,40,60,80 meters
Cluster Probability	Cprob	0.5

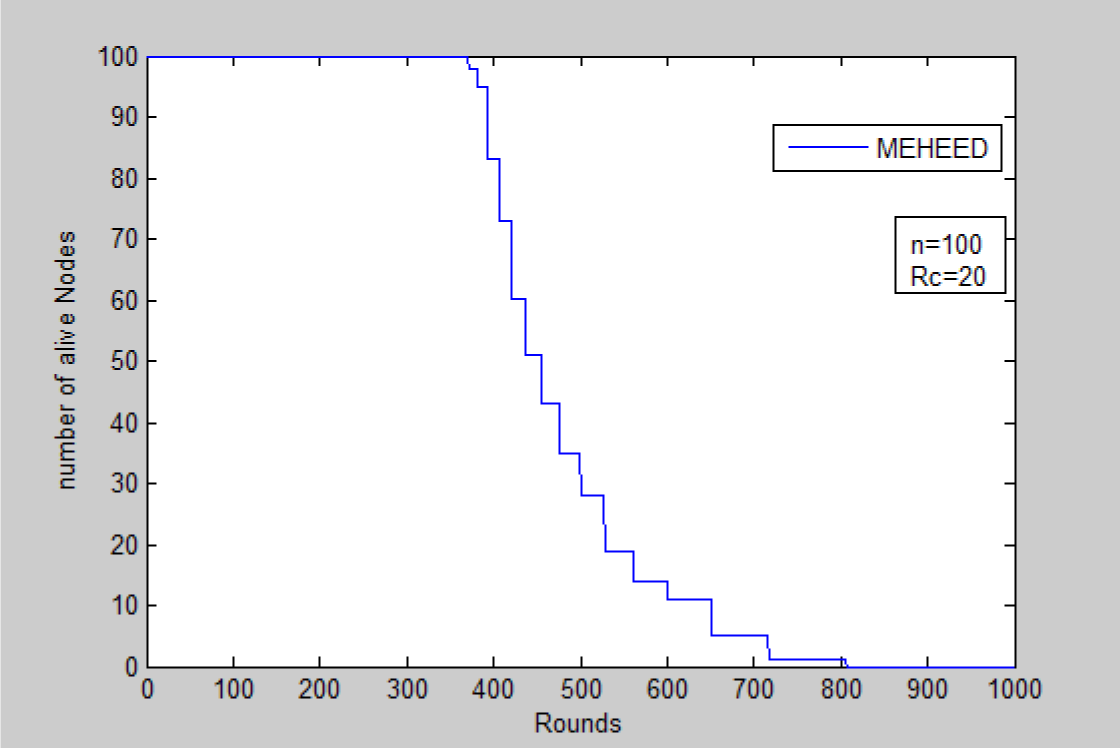


Figure 4.1- Performance of MEHEED

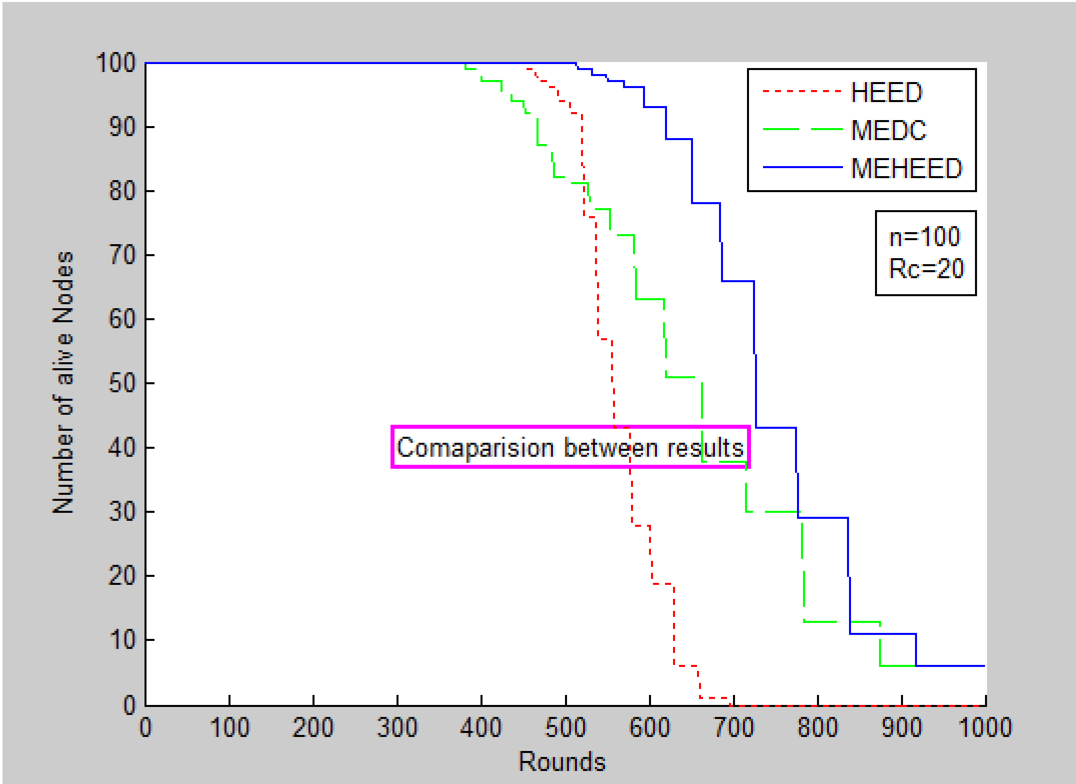


Figure 4.2- Comparative Results HEED, MEDC, MEHEED number of sensor=100 Rc =20

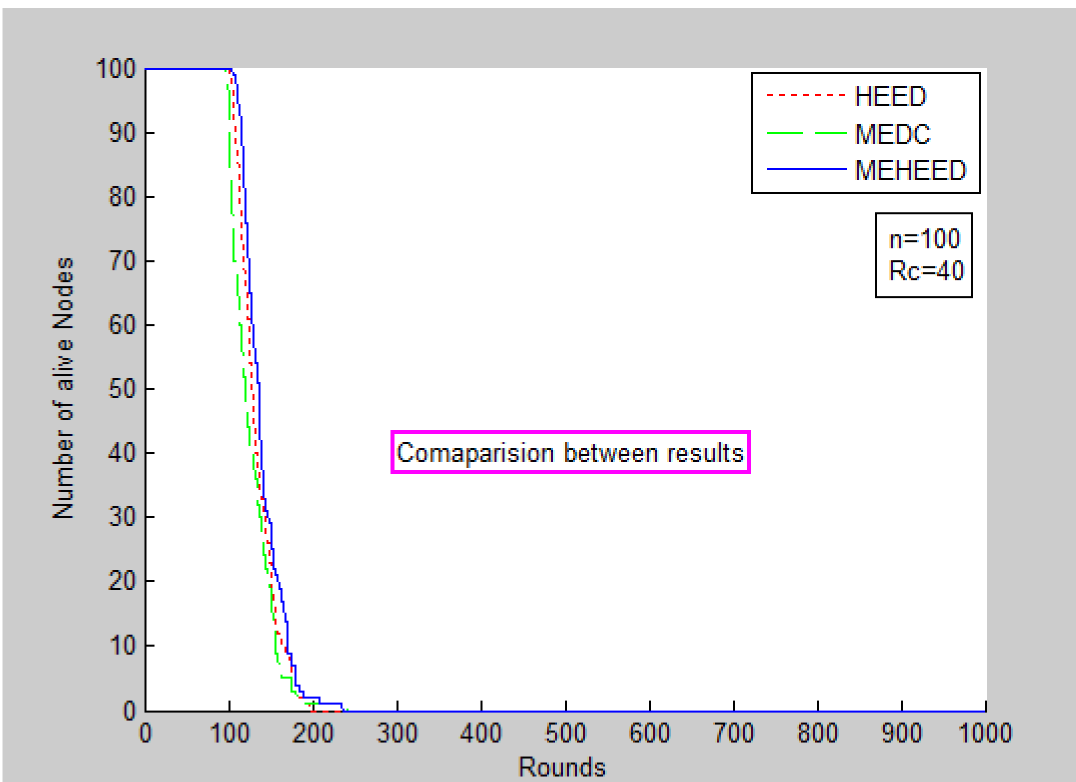


Figure 4.3- Comparative Results HEED, MEDC, MEHEED number of sensor=100 Rc =40

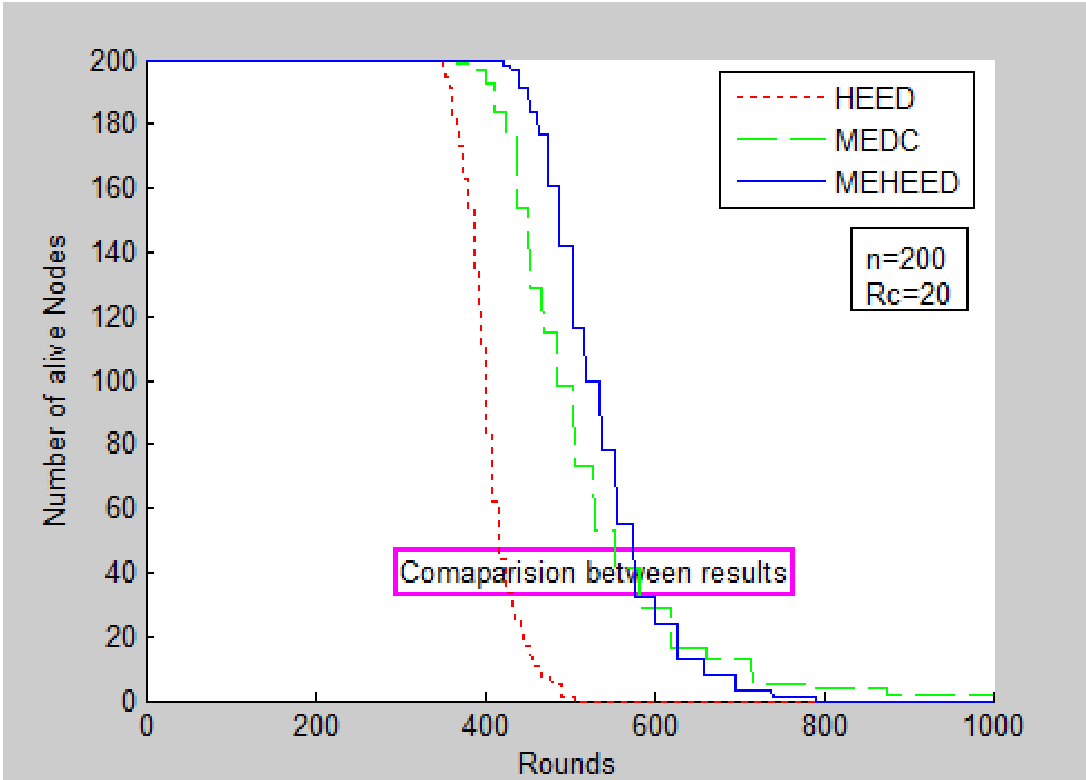


Figure 4.4- Comparative Results HEED, MEDC, MEHEED number of sensor=200 Rc =20

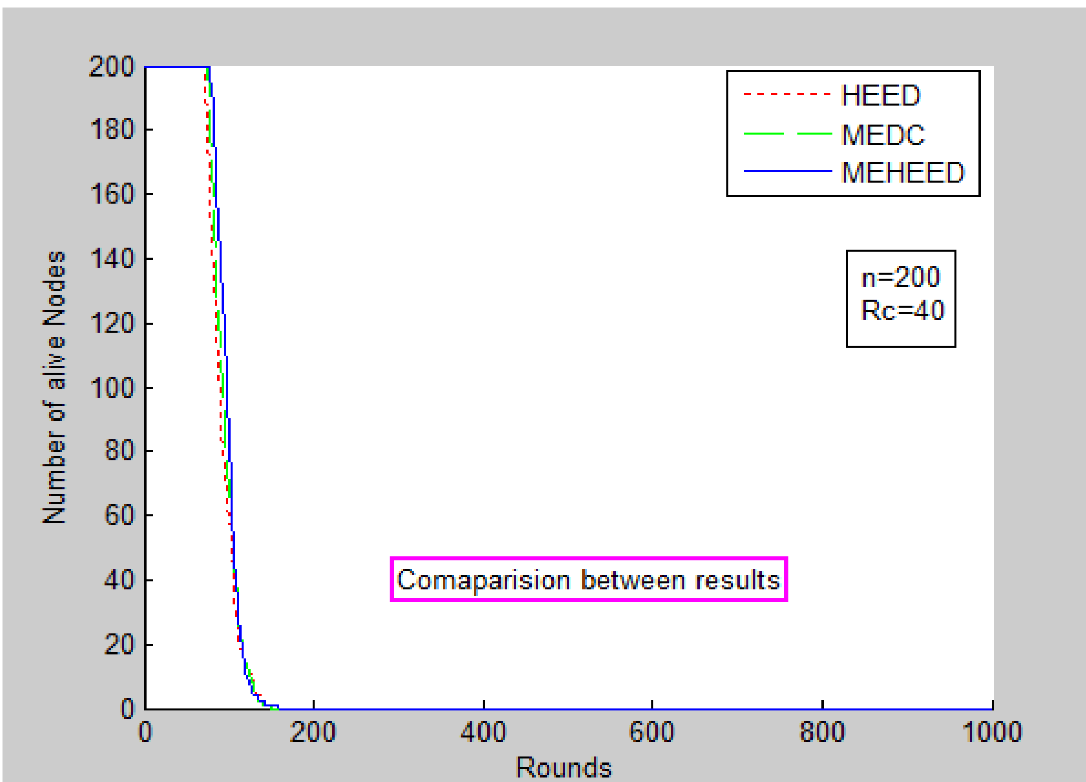


Figure 4.5- Comparative Results HEED, MEDC, MEHEED number of sensor=200 Rc =40

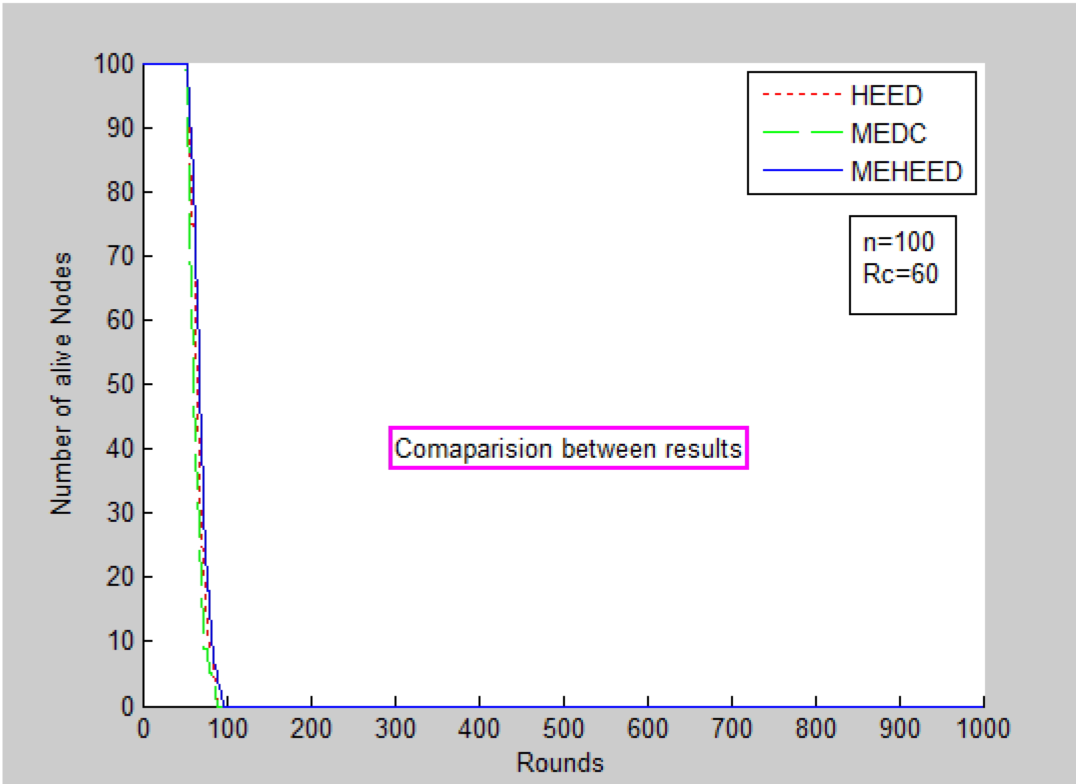


Figure 4.6-Comparative Results HEED, MEDC, MEHEED number of sensor=100 Rc =60

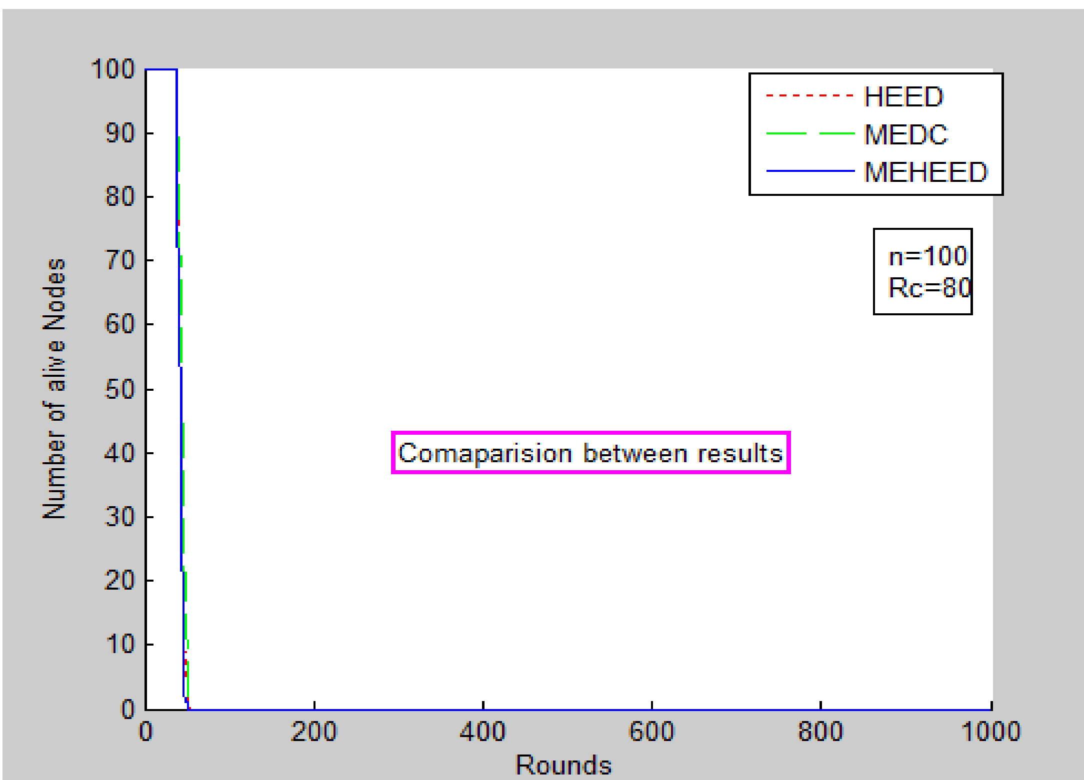


Figure 4.7-Comparative Results HEED, MEDC, MEHEED number of sensor=100 Rc =80

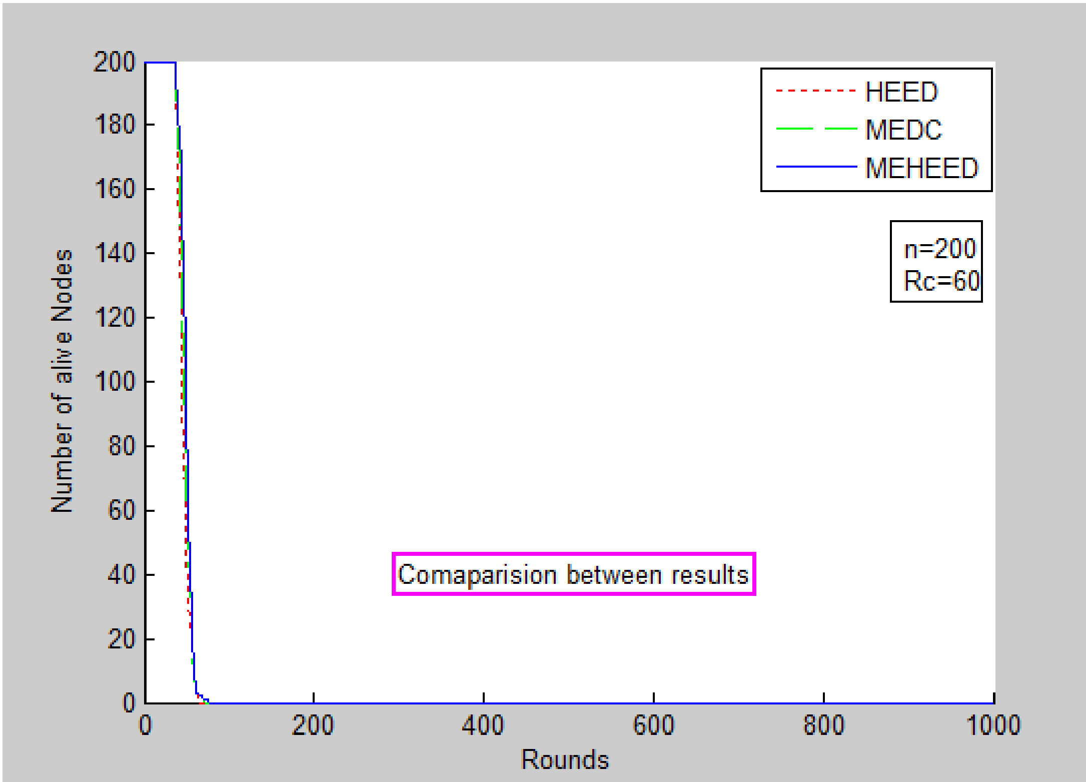


Figure 4.8-Comparative Results HEED, MEDC, MEHEED number of sensor=200 Rc=60

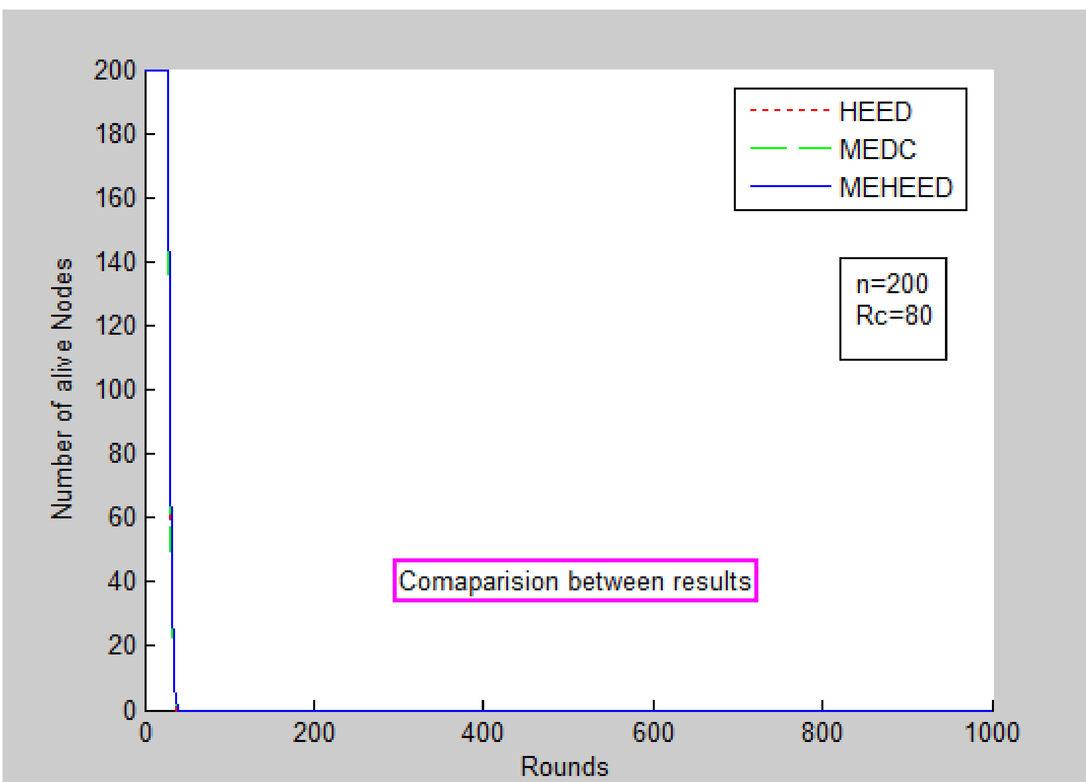


Figure 4.9-Comparative Results HEED, MEDC, MEHEED number of sensor=200 Rc =80

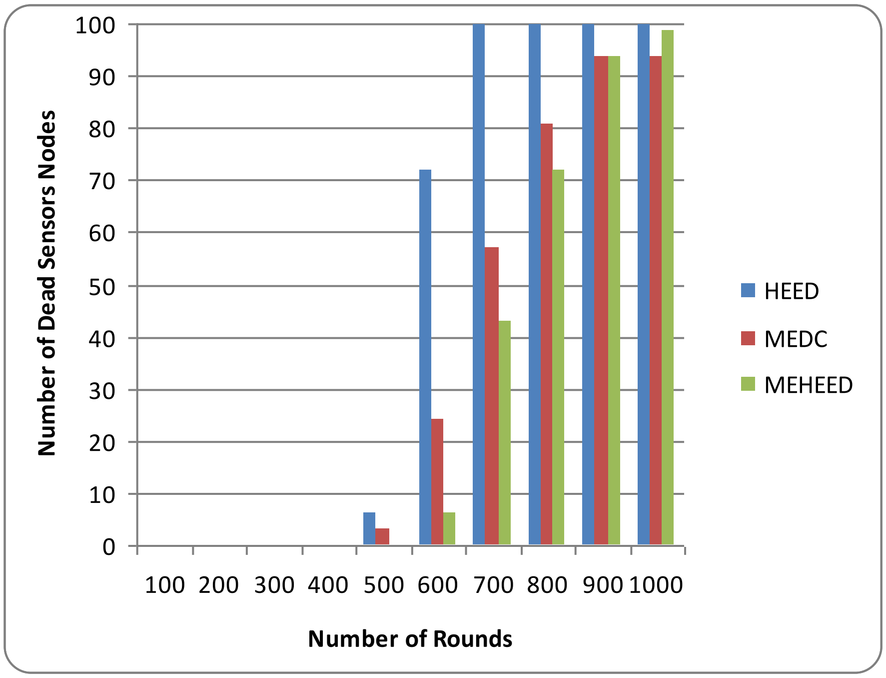
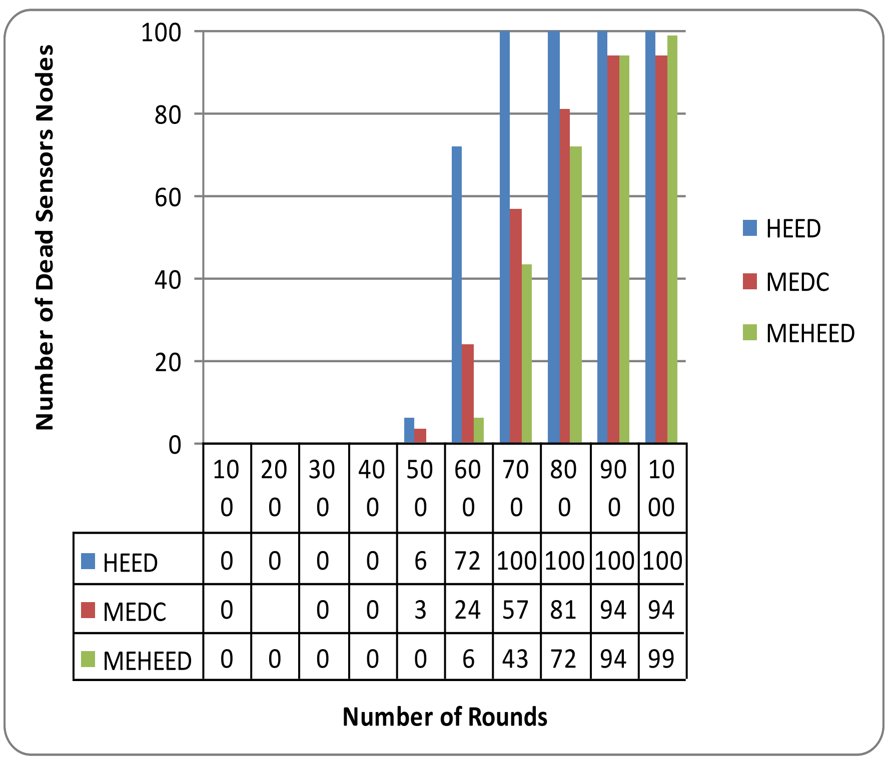


Figure 4.10-Comparison of No of dead sensors V/S Rounds when n=100 and Rc=20.



	100	200	300	400	500	600	700	800	900	1000
HEED	0	0	0	0	6	72	100	100	100	100
MEDC	0		0	0	3	24	57	81	94	94
MEHEED	0	0	0	0	0	6	43	72	94	99

Figure 4.11-Comparison of No of dead sensors V/S Rounds when n=100 and Rc =20 along with detailed table.

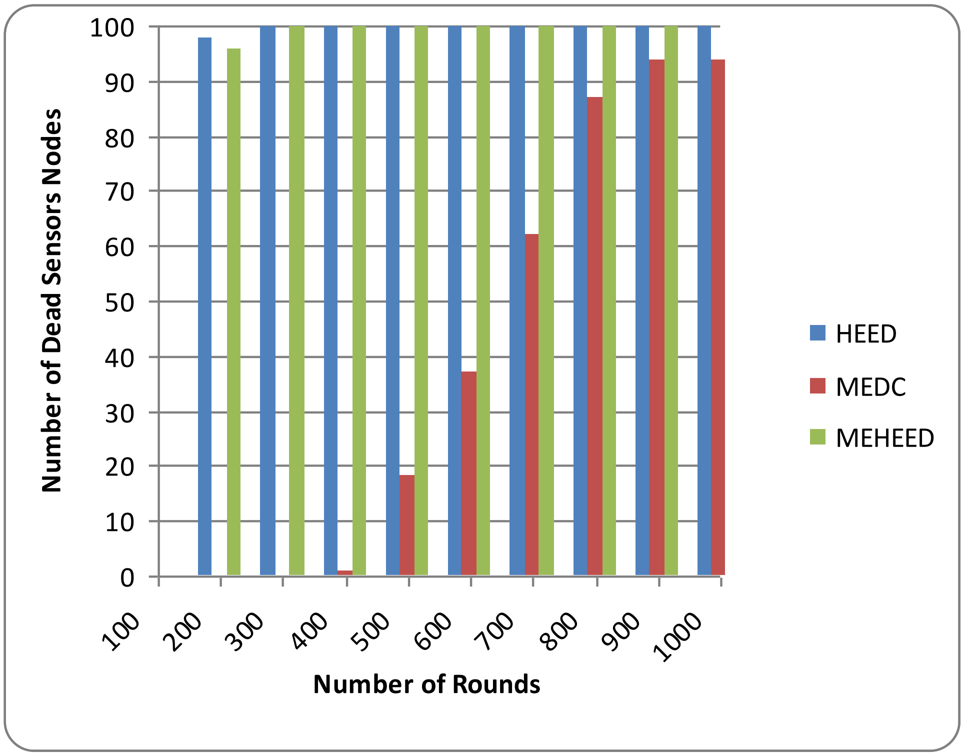
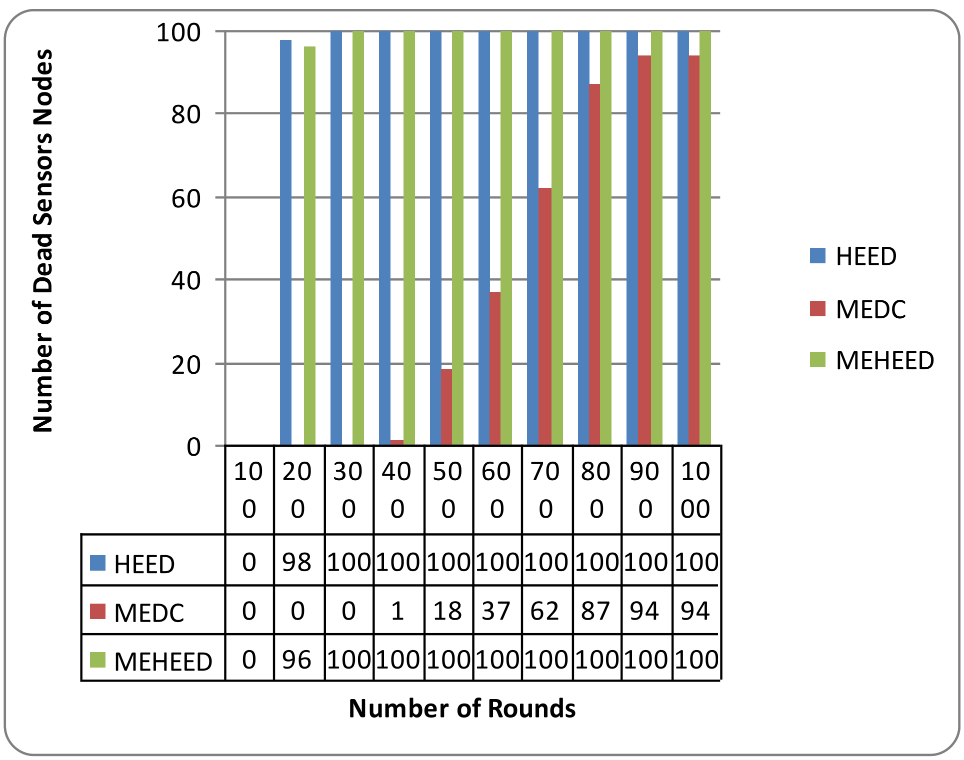


Figure 4.12-Comparison of No of dead sensors V/S Rounds when n=100 and Rc=40.



	100	200	300	400	500	600	700	800	900	1000
HEED	0	98	100	100	100	100	100	100	100	100
MEDC	0	0	0	1	18	37	62	87	94	94
MEHEED	0	96	100	100	100	100	100	100	100	100

Figure 4.13- Comparison of No of dead sensors V/S Rounds when n=100 and Rc =40 along with detailed table

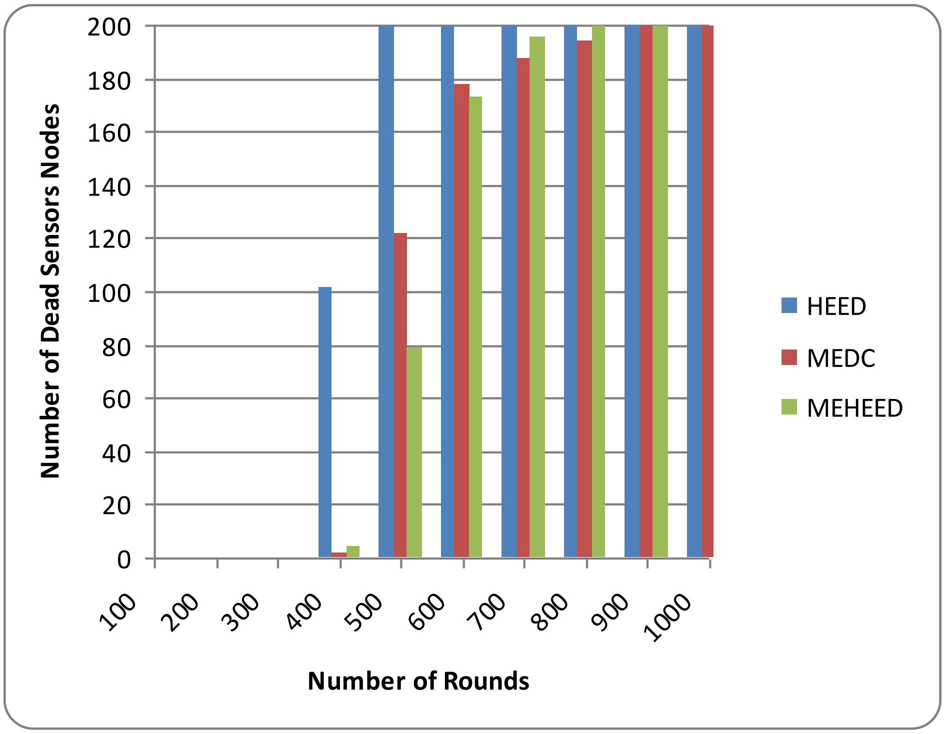
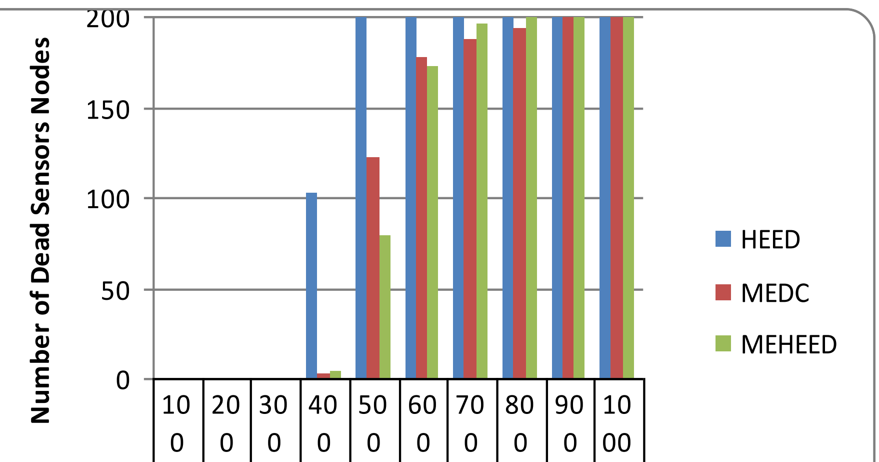


Figure 4.14-Comparison of No of dead sensors V/S Rounds when n=200 and Rc=20.



	100	200	300	400	500	600	700	800	900	1000
HEED	0	0	0	102	200	200	200	200	200	200
MEDC	0	0	0	2	122	178	188	194	200	200
MEHEED	0	0	0	4	79	173	196	200	200	200

Number of Rounds

Figure 4.15-Comparison of No of dead sensors V/S Rounds when n=200 and Rc =20 along with detailed table

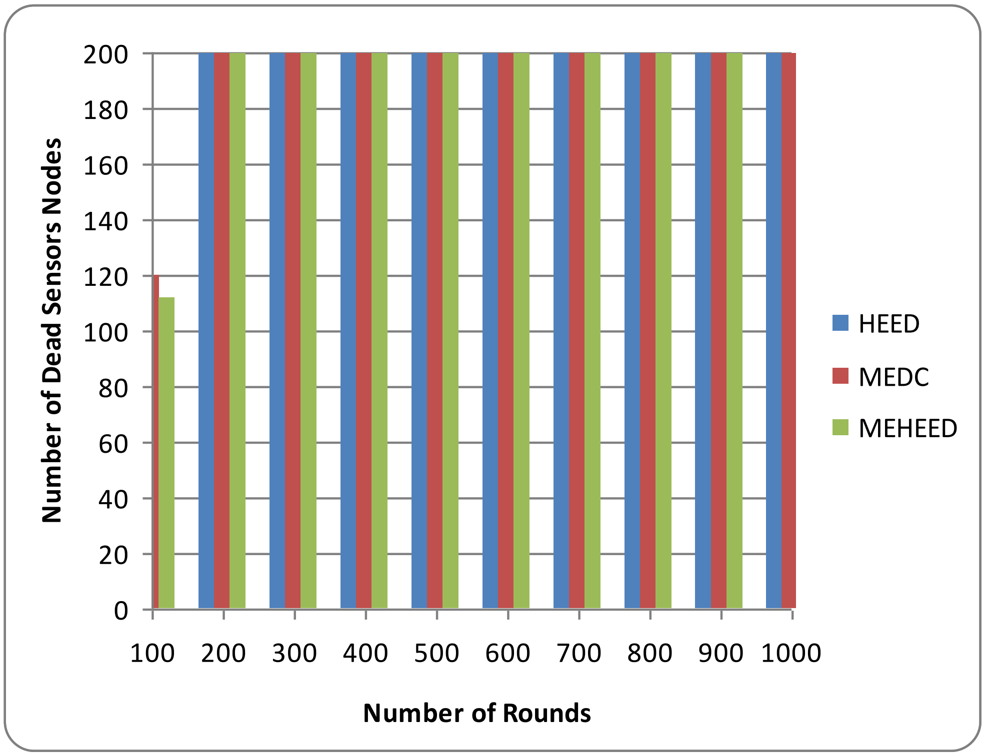
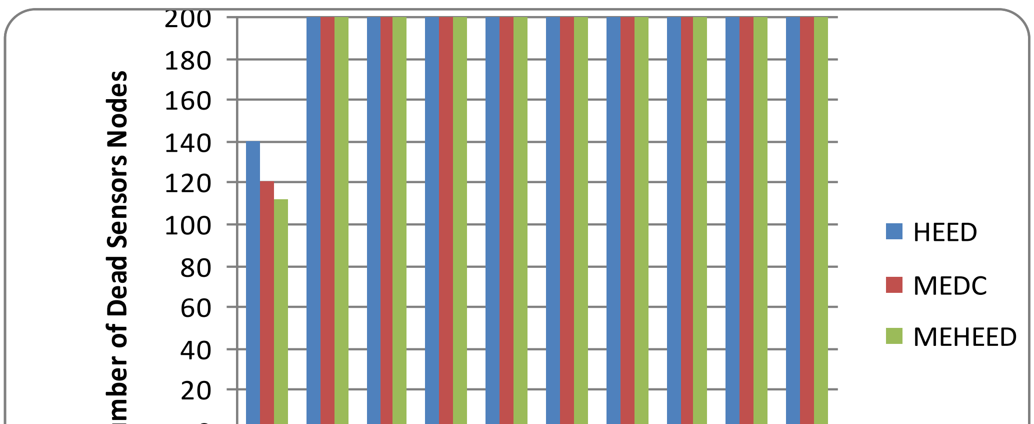


Figure 4.16-Comparison of No of dead sensors V/S Rounds when n=200 and Rc =40.



	100	200	300	400	500	600	700	800	900	100 0
HEED	140	200	200	200	200	200	200	200	200	200
MEDC	120	200	200	200	200	200	200	200	200	200
MEHEED	112	200	200	200	200	200	200	200	200	200

Number of Rounds

Figure 4.17- Comparison of No of dead sensors V/S Rounds when n=200 and Rc =40 along with detailed table

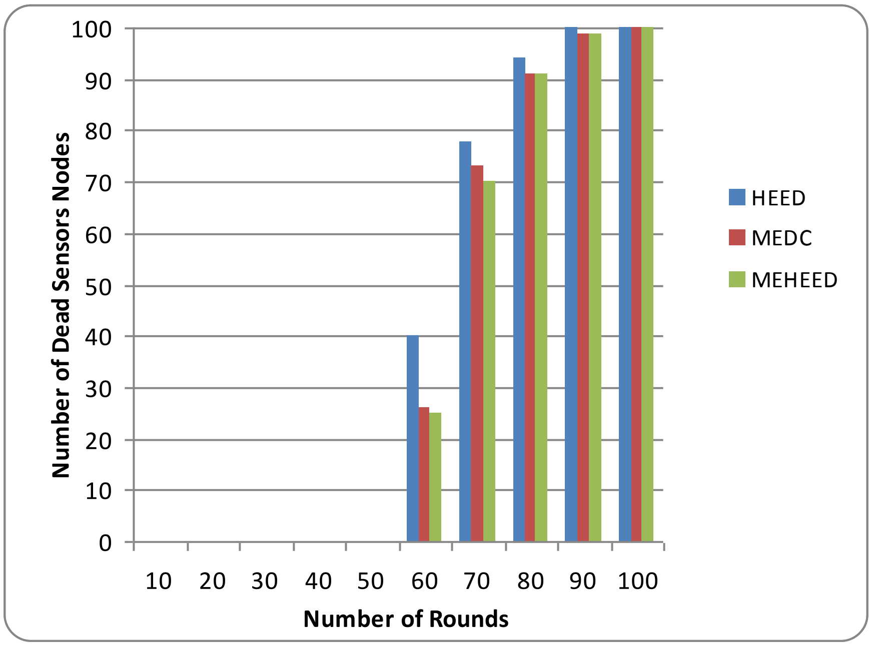
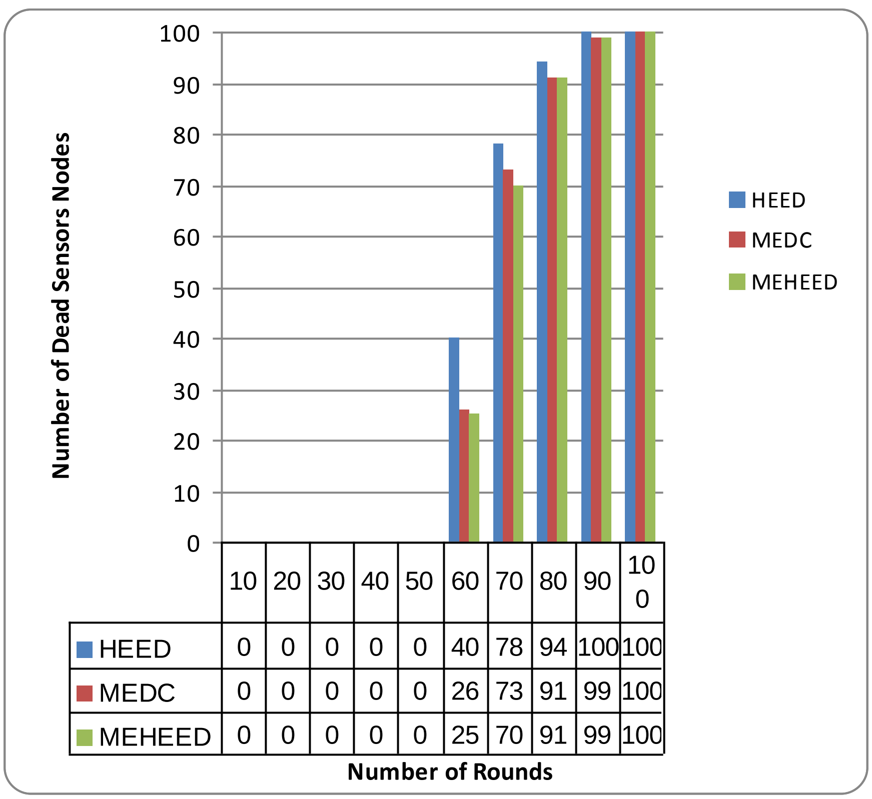


Figure 4.18-Comparison of No of dead sensors V/S Rounds when n=100 and Rc=60



	10	20	30	40	50	60	70	80	90	10 0
HEED	0	0	0	0	0	40	78	94	100	100
MEDC	0	0	0	0	0	26	73	91	99	100
MEHEED	0	0	0	0	0	25	70	91	99	100

Figure 4.19-Comparison of No of dead sensors V/S Rounds when n=100 and Rc =60 along with detailed table

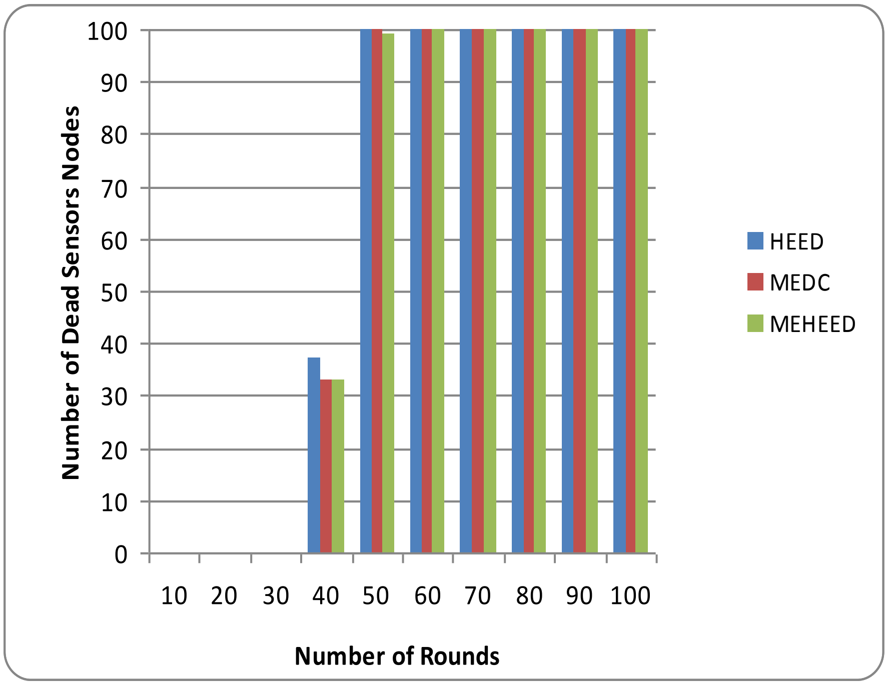
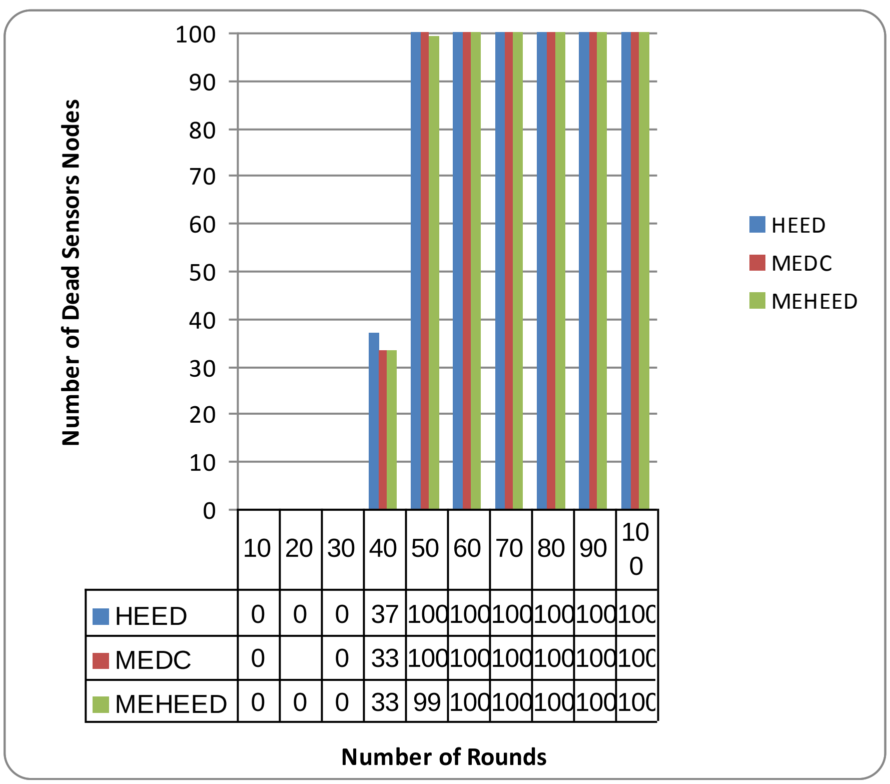


Figure 4.20-Comparison of No of dead sensors V/S Rounds when n=100 and Rc=80



	10	20	30	40	50	60	70	80	90	100
HEED	0	0	0	37	100	100	100	100	100	100
MEDC	0		0	33	100	100	100	100	100	100
MEHEED	0	0	0	33	99	100	100	100	100	100

Figure 4.21-Comparison of No of dead sensors V/S Rounds when n=100 and Rc =80 along with detailed table

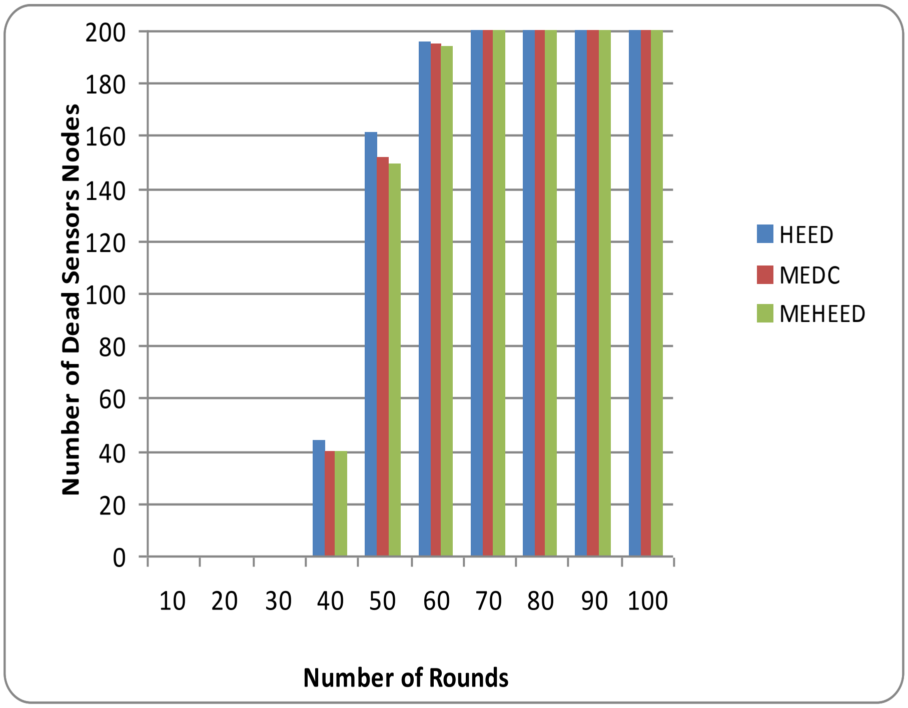
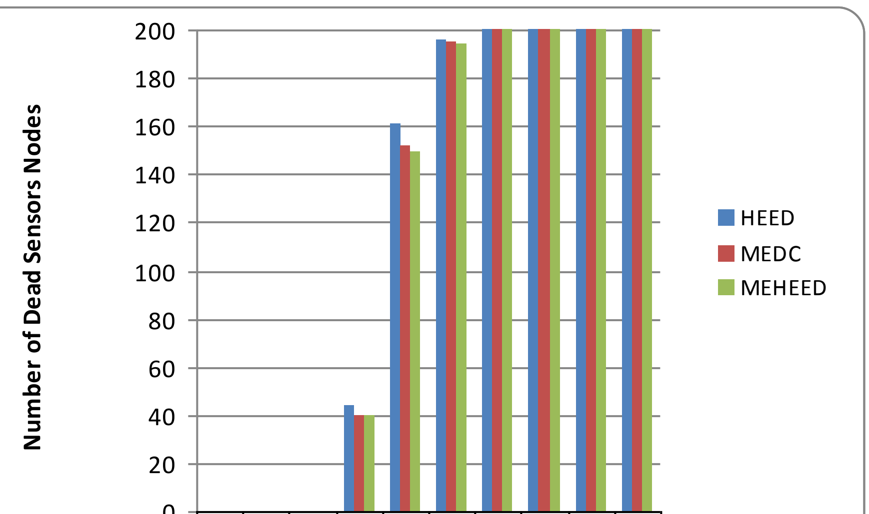


Figure 4.22-Comparison of No of dead sensors V/S Rounds when n=200 and Rc=60



	10	20	30	40	50	60	70	80	90	100
HEED	0	0	0	44	161	196	200	200	200	200
MEDC	0		0	40	152	195	200	200	200	200
MEHEED	0	0	0	40	149	194	200	200	200	200

Number of Rounds

Figure 4.23-Comparison of No of dead sensors V/S Rounds when n=200 and Rc =60 along with detailed table

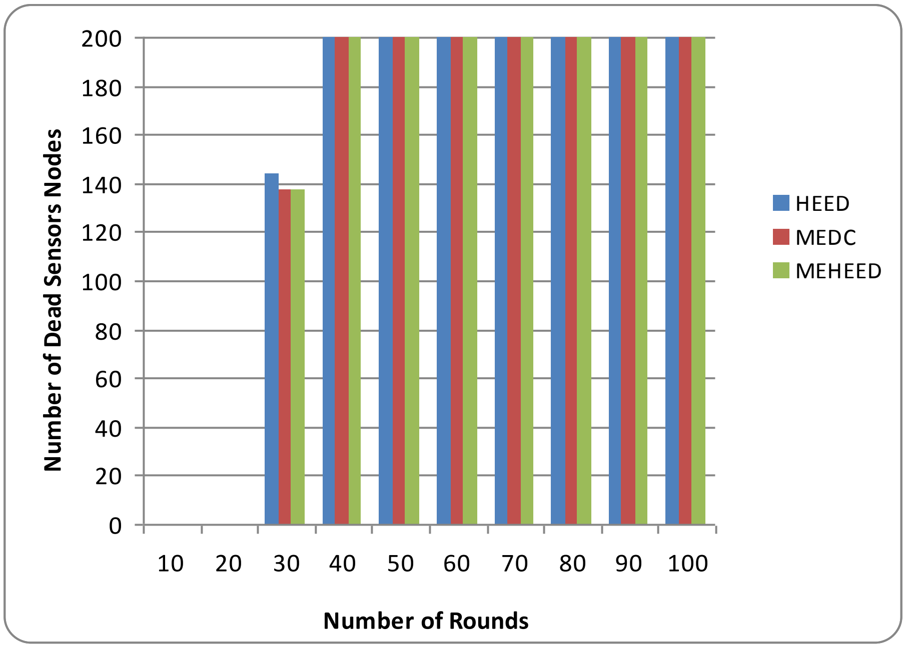
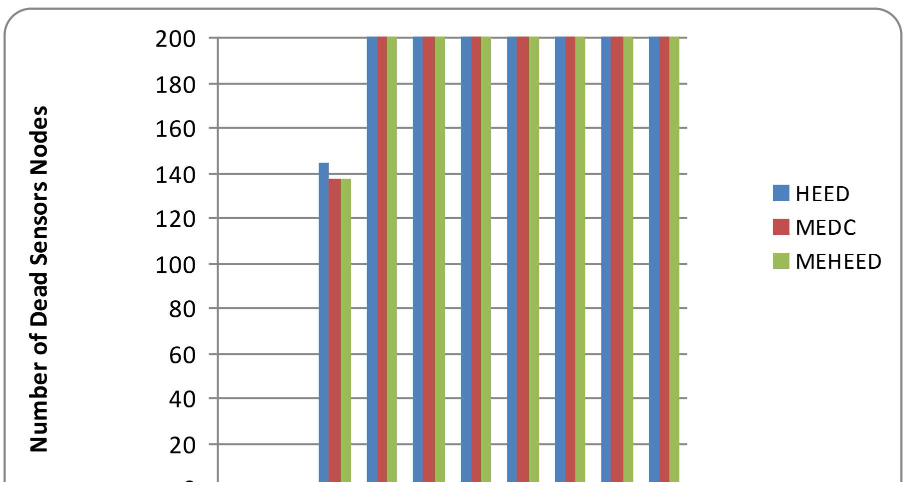


Figure 4.24-Comparison of No of dead sensors V/S Rounds when n=200 and Rc=80



	10	20	30	40	50	60	70	80	90	100
HEED	0	0	144	200	200	200	200	200	200	200
MEDC	0	0	137	200	200	200	200	200	200	200
MEHEED	0	0	137	200	200	200	200	200	200	200

Figure 4.25-Comparison of No of dead sensors V/S Rounds when n=200 and Rc =80 along with detailed table

4.7 Results and Discussions.

MATLAB results for MEHEED are given in section 4.5. Figure 4.1 is experimental result how MEHEED perform. Network life time is measured as number of alive nodes per round. In results graph Y axis represents alive nodes and X axis represents rounds. Result 4.2 is shown for comparative performance of MEHEED, MEDC and HEED on parameter n=100 and range of communication is taken as 20. Red arc represent performance of HEED, Green arc represent MEDC and Blue Arc is representing MEHEED Performance is again measured in terms network life time. Results 4.3-4.9 are again performance analysis of MEHEED, MEDC and HEED by varying values of n and Rc (Range of communication) n is varied on two values 100 and 200. Rc is also varied by four values which are 20, 40, 60 and 80. From different analysis we concluded that MEHEED perform better infect from MEDC also but as parameter are changed graph of MEHEED come closer to MEDC further close to performance of HEED. After n=200 and Rc =40 all three graphs overlap each other. Reason for this decrement in graph of MEHEED; is increased cost of communication. When we increase n from 100 to 200 or Rc from 20 to 40 or to 60 and 80then their will be more no of sensors under range of communication to each other. More sensors under range more advertisements mean more cost of communication which result decrease in performance. Results 4.10- 4.18 are results of these three protocols for dead nodes calculation after rounds. Dead nodes count values after 100 then 200 up to 1000 is taken for all three protocols; and graphs are plotted.

4.8 Conclusion

MEHEED protocol performs clustering and solves network objectives. It enhance network lifetime of network, reduces energy consumption on communication so that network can work up to long time. This protocol has carried advantage of both protocols one is HEED and other is MEDC. MEHEED protocol posses attribute of fast selection along with simplicity. We had worked and proved its efficiency over previous two protocols. MEHEED perform better in most cases even comparable in worst cases. The reason for be betterment is two fold, one is if selection is based on Chprob then computations are skipped a number of steps are overcooked, on the second hand if selection cannot be taken on probability basis then message passing as like MEDC protocol will select cluster heads in simple way. Message passing will be done and under a range of communication sensor with maximum residual energy will be chased as cluster head. Selected cluster head will transmit cluster head declaration message to sensors under its range of communication. Sensor that comes under range of communication will act as cluster members. Cluster head will aggregate sensory information from cluster members and transmit that data to base station.

CHAPTER 5 CHANGE POINT DETECTION USING FUZZY INFERENCE SYSTEM

5.1 Introduction

This work has focused on objective to make network more application specific and to make network more usable. Change point detection is having a number of applications in WSN. Change point detection is different from regular data gathering. This work is carried out to present a merge solution that perform clustering and change point detection together.

Up to know no clustering protocol has focused on this application. We have tried to present change point detection in the midst of MEDC clustering protocol. For change point detection we have worked with fuzzy logic. Section 5.2.1 introduces background of fuzzy logic. Section proceeds with advantages of fuzzy logic and its applications in change point detection. This section also includes some literature survey wherever fuzzy logic is applied in change point detection. After that 5.2.2 section is about MATLAB tool Fuzzy Inference systems (FIS). Section 5.3 present Fire Detection Scenario. Section 5.4 present modified MEDC algorithm along with change point detection solution. Section 5.5 is snapshot of experimental evaluation on FIS. Section 5.6 will be discussions about results after that section 5.7 will conclude this work.

5.2 Background

5.2.1 Fuzzy Logic

Fuzzy logic was introduced in 1965 to handle vague concepts or can say partial truth. Fuzzy logic has represented on multi valued logics. Fuzzy logic can also present partial truth values which may range from a [0, 1] in contrast of traditional systems which represent the fact as completely false or completely true. In fuzzy logic theory truth value of any statement is represented in degree. Degree of truth maps the statement's belongings to a particular set. Fuzzy logic has multidisciplinary nature and fuzzy inference systems are termed with many names example includes fuzzy-rule-based systems some may call fuzzy expert systems, or fuzzy modeling. A number of advantages are associated with fuzzy logic which is as follows

Advantages of fuzzy Logic

- The biggest advantage of fuzzy logic is it can also represent linguistic variables.
- Fuzzy logic concepts are easy to understand and to implement also. Fuzzy logic includes minimum complexity and can be implemented in natural ways.
- Fuzzy logic is flexible enough that it can work with imprecise data.
- Fuzzy logic is also capable to represent non linear function without increasing complexity level to excess.
- Fuzzy logic is not depended on dense model or training data. Design experts can design on basis of their experiences.
- Fuzzy logic can be combined with Neural networks and result Adaptive Neuro-Fuzzy theory
- Fuzzy logic can provide solution of many complex problems exists in different fields like fields of medicine
- Fuzzy logic has resemblance with human reasoning and also with decision making so it may also answers all type of uncertainties and ambiguities which was near to impossible with discrete terms.
- Fuzzy is useful method for change point detection method as it starts with imprecise data, build fuzzy sets and fuzzyfied them from probability function for concluding result from fuzzy rules.
- Fuzzy systems also have their applicability in various fields like expert systems, automatic control, decision analysis, data classification and computer vision.

In 1999 author Metternicht has used same approach for computation of changes occurred in remotely sensed area for map revision [Metternicht, et al., 1999] The author has formulated fuzzy sets and fuzzy rules for defining the likelihood of changes detected from remotely sensed data then after they used decision support system to conclude change. Authors had shown that fuzzy logic has given better change detection results in contrast of traditional systems like image ratio etc. results are shown to prove how change point detection using fuzzy theory help to conclude map revision more accurately. In paper [Zlokolica, et al., 2005] author had worked on application concerned on tracking and denoising of image sequence. The work has contributed fuzzy recursive motion detector. The author has worked by inputting noisy sequence with fuzzy logic motion detector in order to determine the degree of motion confidence. They proposed motion detection algorithm; concentrating on two criteria one is robustness to noise and second is changing illumination conditions and motion blur in temporal recursive denoising. Further in 2011 the author D'Angelo used the method of fuzzy logic to find faults in machines example of such machine is; induction machine stator-winding [D'Angelo, et al. 2011]. The author has used the approach of change point detection for improved forecast of the new failure detection procedure against false alarms, combined with a good sensitivity that allows the detection of rather small fault signals. They had defined and implemented two steps procedure. First step is transformation of initial data using fuzzy clustering. Second is change point detection with the help of Metropolis–Hastings algorithm.

5.2.2 FIS

FIS is tool of MATLAB that interprets input values on basis of designed rules and produces output. Working of FIS include five steps

- Fuzzify Inputs
- Apply Fuzzy Operator
- Apply Implication Method
- Aggregate All Outputs
- De-fuzzify

a) Fuzzify Inputs

This is first step which take the inputs and calculate the degree to which they belong to each of the suitable fuzzy sets with the help of membership functions. We take inputs always in crisp form after which Fuzzy Logic Toolbox outputs a fuzzy degree of membership in the qualifying linguistic set.

b) Apply Fuzzy Operator

In Second step FIS apply operators. Two main operators are available in FIS one is AND another is OR. Fuzzyfied inputs are matched to antecedents of rules. If the antecedent of a given rule has more than one part, the fuzzy operator is applied to obtain number which represents the result of the antecedent for that rule. This number will then be applied to the output function. The input to the fuzzy operator is two or more membership values from fuzzyfied input variables. The output is a single truth value.

c) Apply Implication Method

After weighting assigned to each rule, the implication method is implemented. The input for the implication process is a single number given by the antecedent, and the output is a fuzzy set. Implication is implemented for each rule. Two built-in methods are supported, and they are the same functions.

d) Aggregate All Outputs

In Aggregation step fuzzy sets that are representing the outputs of each rule they are combined into a single fuzzy set. Decisions are based on the testing of all of the rules in an FIS; the rules must be combined in some manner in order to make a decision. Aggregation only occurs once for each output variable.

e) De-fuzzify

The input of Defuzzification step is a fuzzy set which is received from previous step (the aggregated output fuzzy set). Output of Defuzzification is a single number. However, the aggregate of a fuzzy set encompasses a range of output values, and so must be defuzzified in order to resolve a single output value from the set.

Types of FIS

There are two types of FIS in MATLAB.

- Mamdani
- Sugeno

a) Mamdani

Maamdani FIS is most commonly used FIS method developed by Ebrahim Mamdani in 1975. The development was aimed to control some mechanical devices with the help of rules. The first idea was based on deriving conclusion from rules designed by experts. Now days the inference process has taken new shape yet the basic idea is same as proposed by authors in research paper from Ebrahim Mamdani. This tool expects that the output membership functions should be from fuzzy sets. It may also possible to represent output function as a single output function instead of sets. At Defuzzification step efficiency could be increased using Pre-Defuzzified fuzzy sets because it simplifies computations. Mamdani FIS is intuitive in nature having acceptance world wide for human given inputs.

b) Sugeno

Sugeno FIS was proposed by author Takagi-Sugeno-Kang in 1985. Sugeno FIS works mainly for mathematical functions. Sugeno FIS can also work with linear, adaptive and optimization techniques.

Selection of Mamdani or Sugeno FIS is done as shown in Figure 5.1

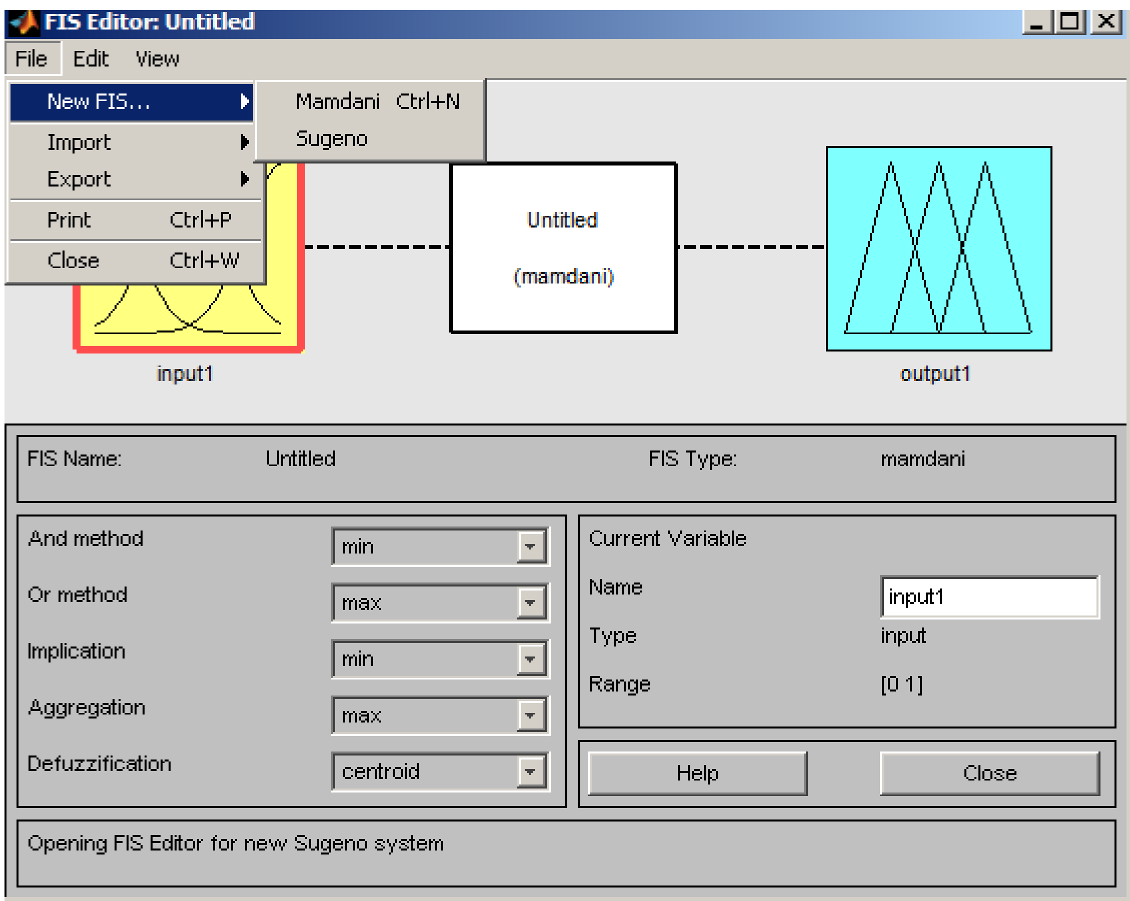


Figure 5.1- Selecting FIS Mamdani or Sugeno

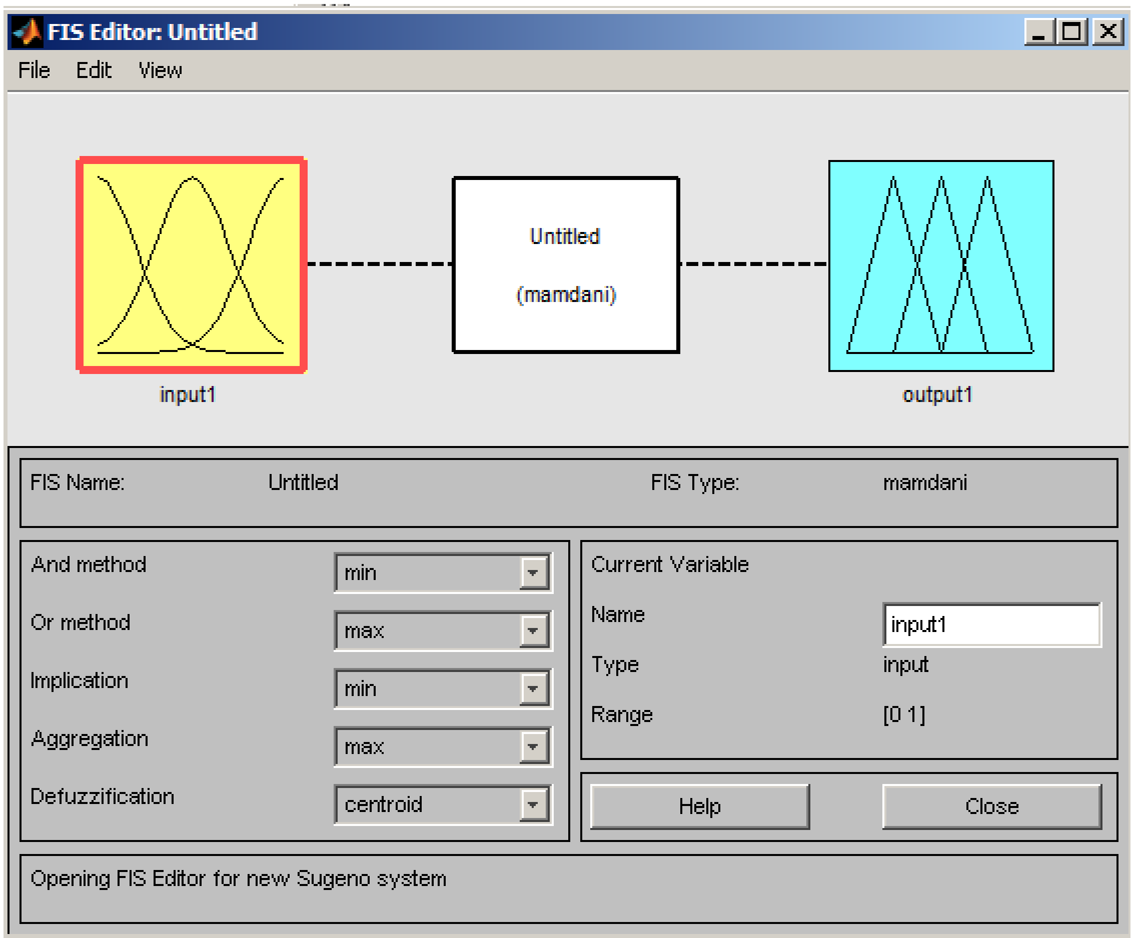


Figure 5.2- First View of Mamdani FIS

The first window of Mamdani FIS appears is shown in figure 5.2. The input and output variables can be added by selecting edit option next click on variables as shown in Figure 5.3

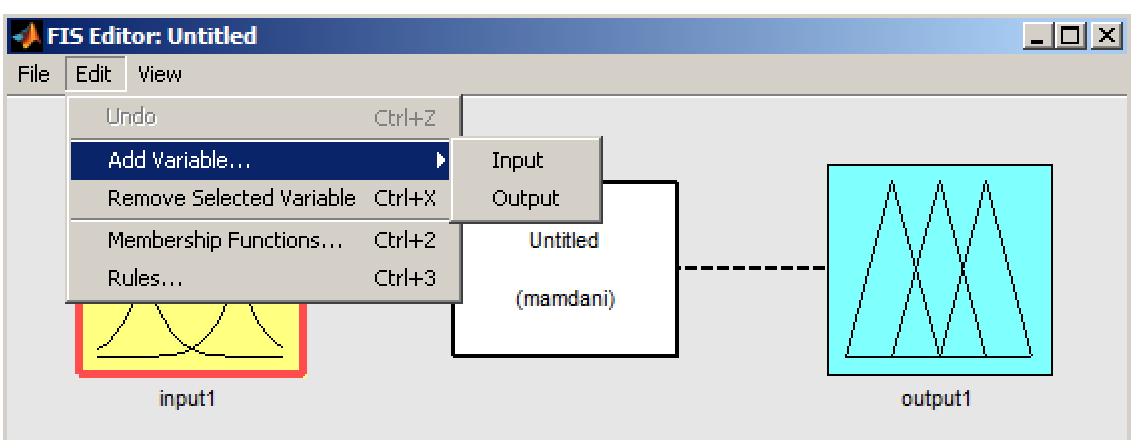


Figure 5.3 Adding input and Output Variables

It is also possible to remove variables by selecting next option. Variable can have multiple membership function. Membership function can be named under name bar and its type can be selected from drop down list shown in figure5.4. The Range and Display range parameters are value representing on x axis. Under a Variable membership functions could be added or deleted by selecting edit option as shown in figure 5.5. Under edit there is one option named as Rule. Here we defined rules as decided by experts. Next option is View under view we can have two options one is rules and other is surface. On clicking rule tab rule editor will open as shown in figure 5.7. Rule editor maps the rule between input variables and output variables using connections

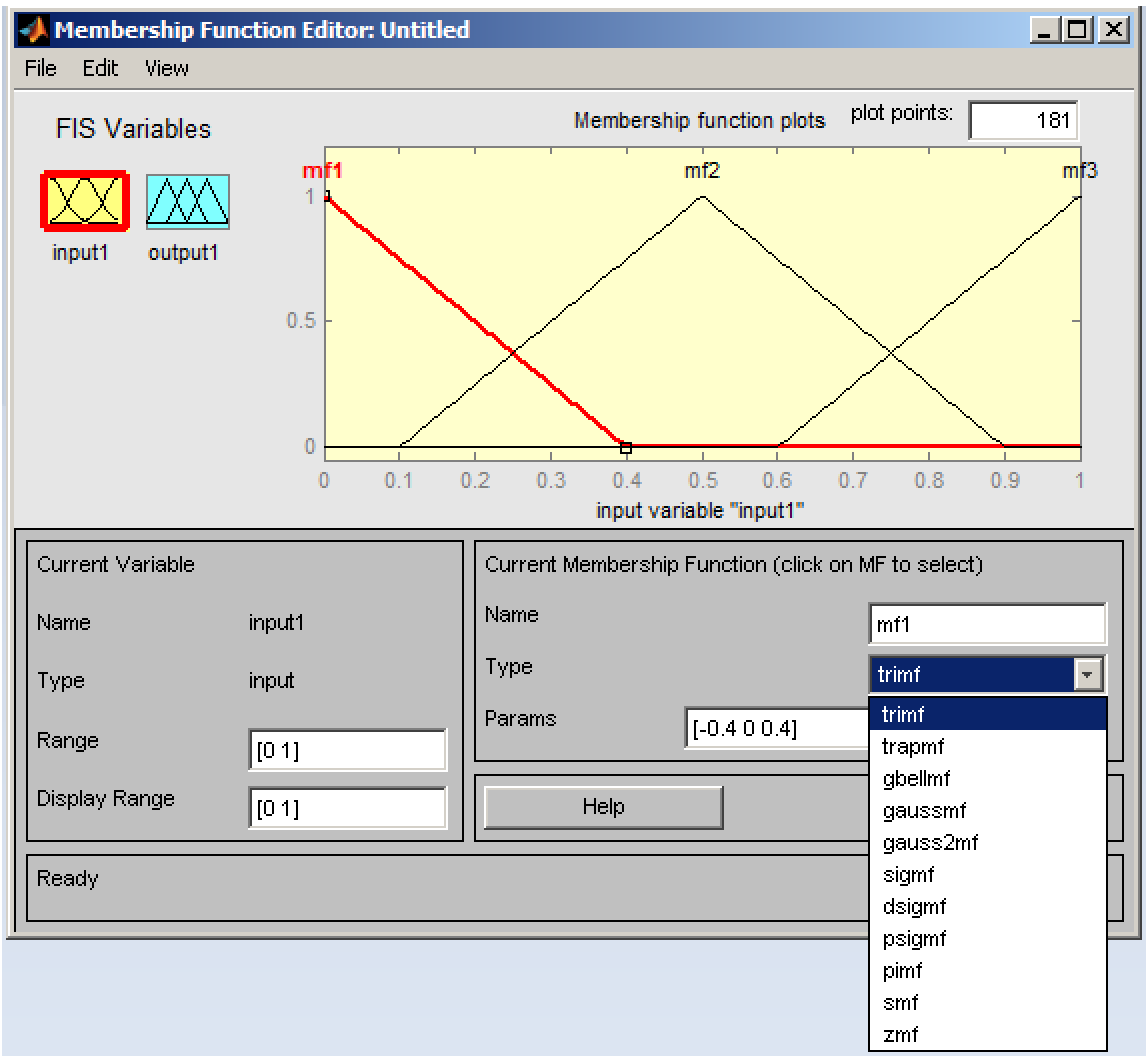


Figure 5.4 Selecting Membership functions

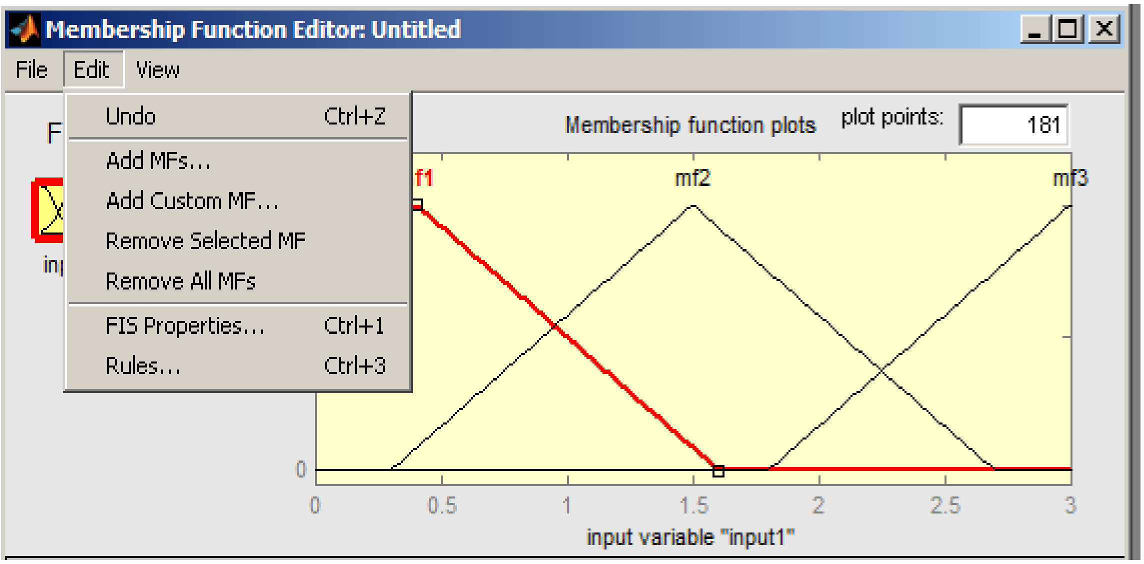


Figure 5.5- Adding Membership Functions

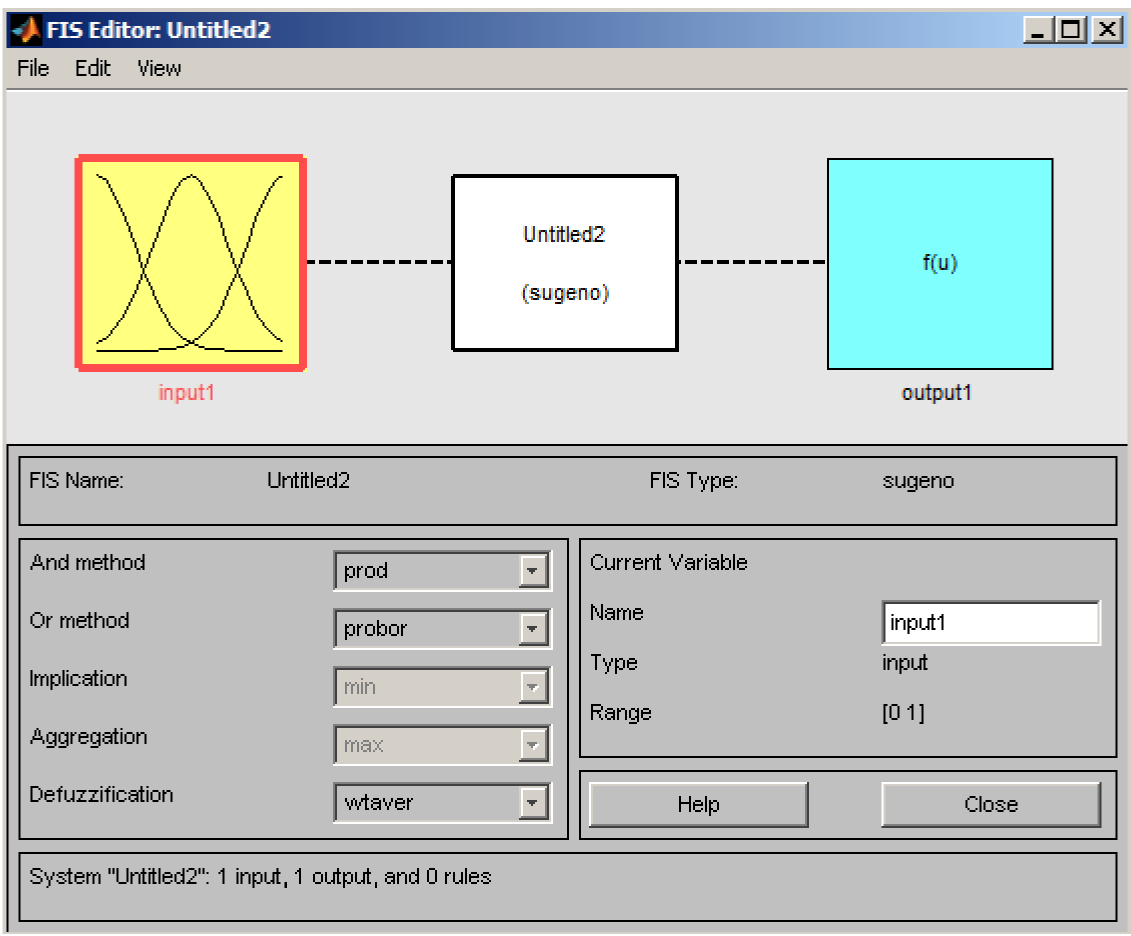


Figure 5.6 - First View of Sugeno FIS

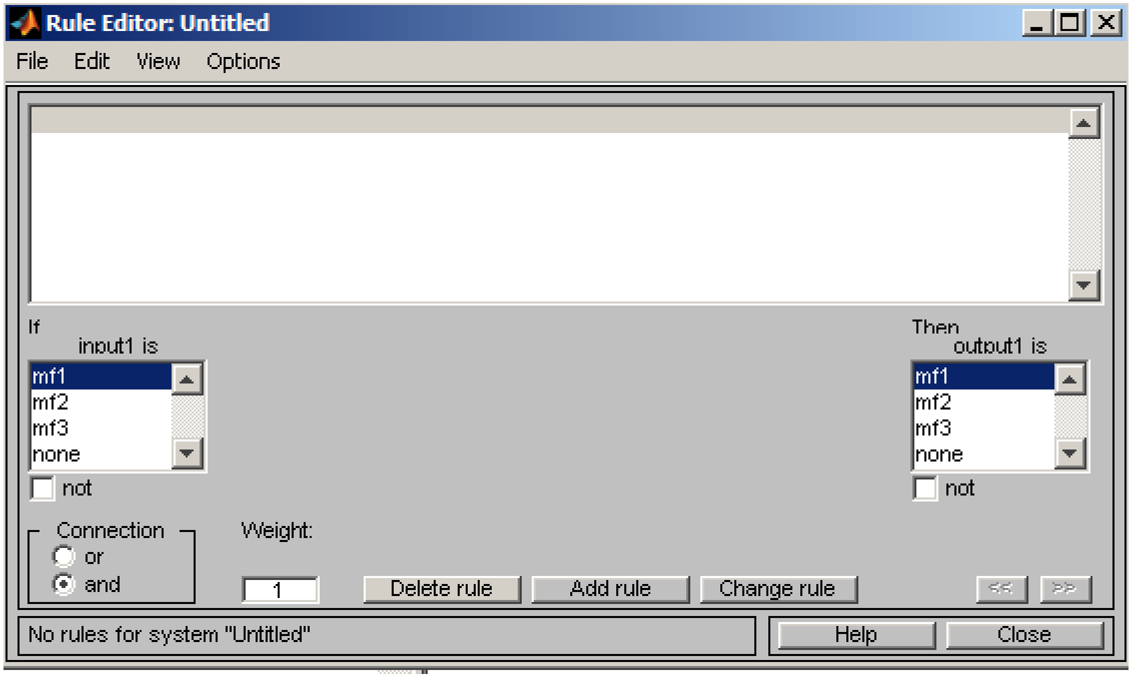


Figure 5.7- Rule editor

5.3 Fire Detection Scenario

Fire detection is application of change point detection. In case of fire environmental parameters have those values which deviate from parameters under normal situation. Here for fire detection scenario we are considering that sensor are motes who are sensing three parameters one is Heat index, second is Relative Humidity and third parameter is Carbon Monoxide. We are assuming that sensors will sense the values and send sensory values to their cluster heads which are chased by MEDC clustering protocol. Cluster heads will aggregate the received data and for aggregation we are taking simple averaging rule. These three aggregated values will be inputted into fuzzy system of cluster heads. Fuzzy system will decide on basis of fuzzy rules and decide whether these values are concluding presence of fire.

Three Input Variables

- Heat index
- Relative Humidity
- Carbon Mono-Oxide

One Output Variables- Fire probability

5.4 MEDC with Change Point Detection Algorithm

Part 1 Procedure Cluster_formation (n)

1. For each IDi
2. Counter =0
3. For each IDj within Rf of IDi
4. Advertise Eresidual_i
5. For each IDi
6. Put all incoming advertisements from sensors j into Qi
7. For each IDi
8. While Qi is not empty
9 If Eresidual_i <= Eresidual_j
10. Send ok message to IDj
11. Counter = 1
12. Else
13. Delete this advertisement from queue
14. For each IDi
15. If counter =0
16. Send cluster_head_declaration message to IDj within Rf

Part 2 Procedure Change_point_detection (n)

1. For each IDi
2. If counter =0
3. Collect the information from sensors IDj from which OK message received.
4. Calculate mean of sensed information.
5. Input the average values in FIS.
6. Apply fuzzy rules and calculate Pr
7. If Pr >=50
8. Report to base station for change point detection.

5.5 Experimental Evaluation

Figure 5.8 is snapshot of FIS tool. We are taking Mamdani FIS with three input variables and one output variables.

Table5.1- Heat Index with risk level.

Heat Index	Risk level
Less than 91°F	Lower
91° to 103°F	Moderate
103° to 115°F	High
Greater than 115°F	Very High to Extreme

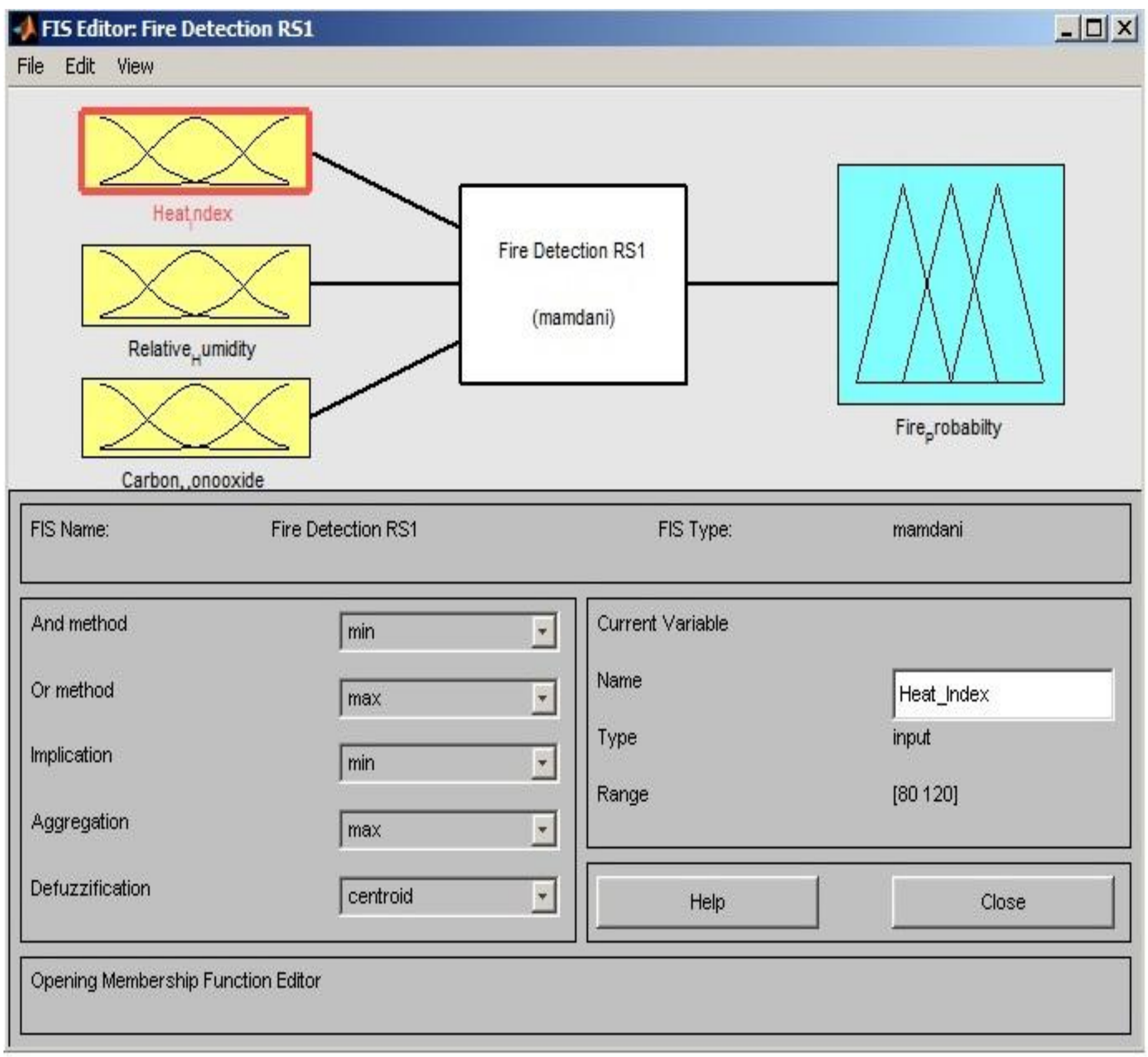


Figure 5.8- MATLAB Implementation Snapshot for fire detection

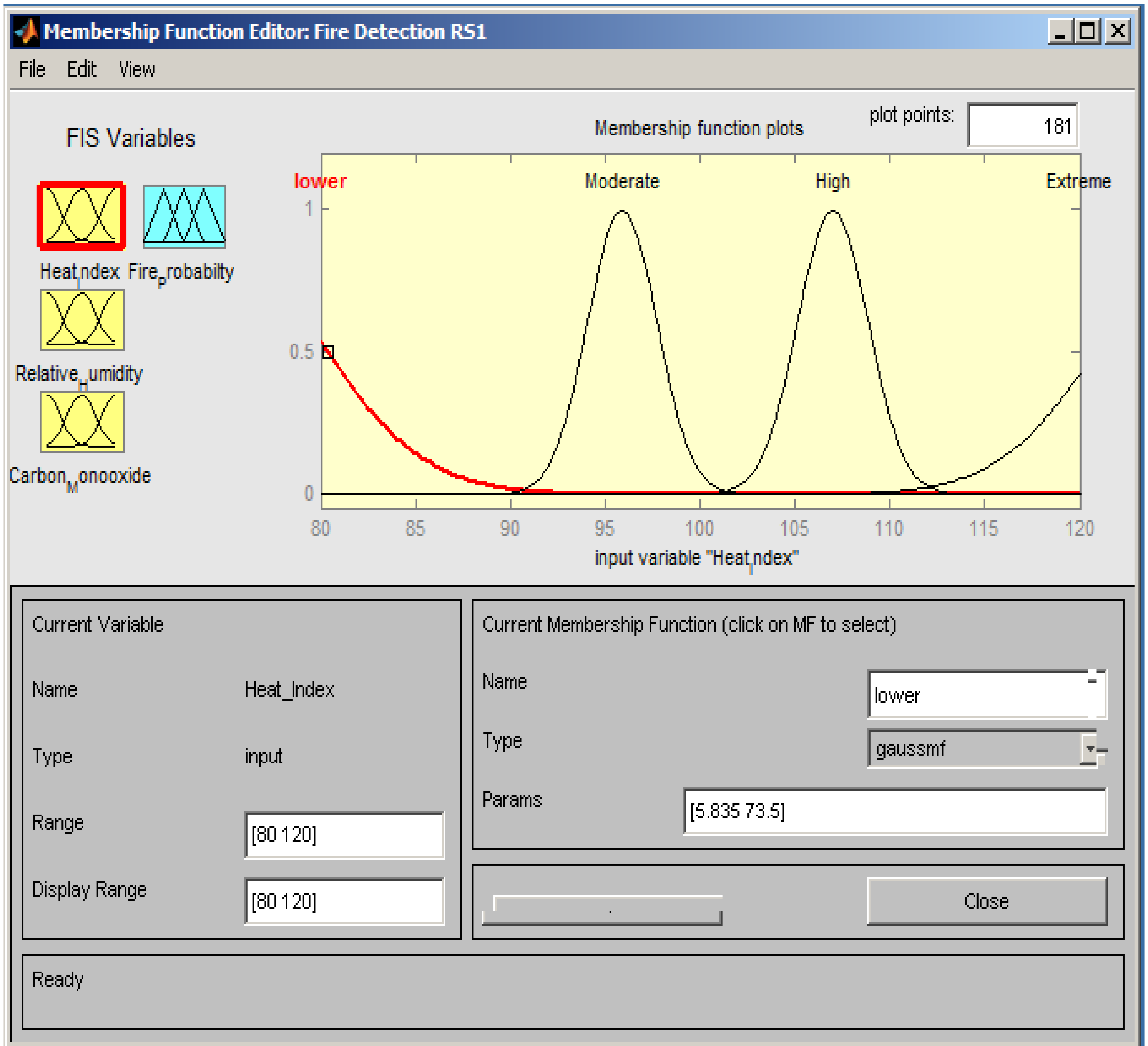


Figure 5.9- Snapshot for input variable Heat Index

Figure 5.9 is snapshot of membership function plot over range [80 120] for Heat index input variable. We have taken four membership functions (mf) lower, moderate, high and extreme. This data is taken from Occupational Safety and health administrations refer Table 5.1Figure 5.10 is shown snapshot for input variable Relative Humidity. We have defined four mfs extreme danger, danger, extreme caution and caution. The range value of this parameter is taken from data national weather service heat index. In figure 5.11 we had shown data from national weather services. Data has shown mapping of heat index with relative humidity. From their; we had mapped scale value of relative humidity is from 40 to 100. Relative humidity is measured according to heat index.

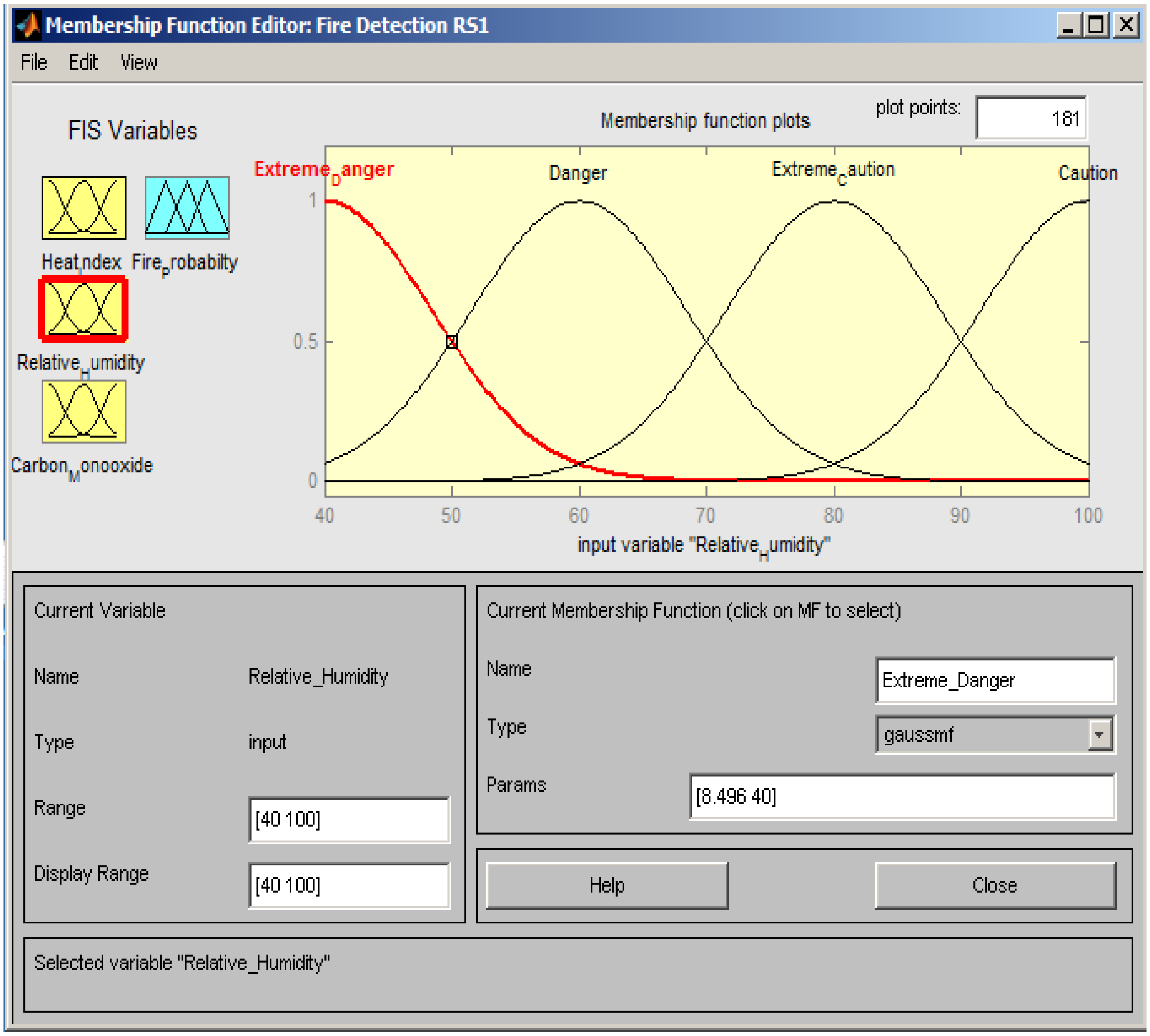
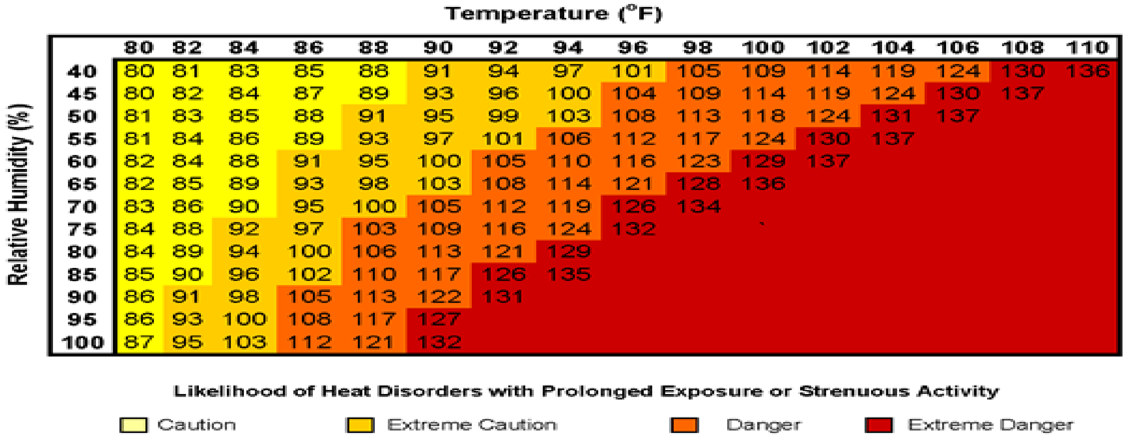


Figure 5.10- Snapshot for input variable Relative Humidity

NOAA's National Weather Service

Heat Index

Temperature (°F)



Relative Humidity (%)	80	82	84	86	88	90	92	94	96	98	100	102	104	106	108	110
40	80	81	83	85	88	91	94	97	101	105	109	114	119	124	130	136
45	80	82	84	87	89	93	96	100	104	109	114	119	124	130	137	
50	81	83	85	88	91	95	99	103	108	113	118	124	131	137		
55	81	84	86	89	93	97	101	106	112	117	124	130	137			
60	82	84	88	91	95	100	105	110	116	123	129	137				
65	82	85	89	93	98	103	108	114	121	128	136					
70	83	86	90	95	100	105	112	119	126	134						
75	84	88	92	97	103	109	116	124	132							
80	84	89	94	100	106	113	121	129								
85	85	90	96	102	110	117	126	135								
90	86	91	98	105	113	122	131									
95	86	93	100	108	117	127										
100	87	95	103	112	121	132										

Likelihood of Heat Disorders with Prolonged Exposure or Strenuous Activity

Caution | Extreme Caution | Danger | Extreme Danger

Figure 5.11 Heat index with Relative humidity

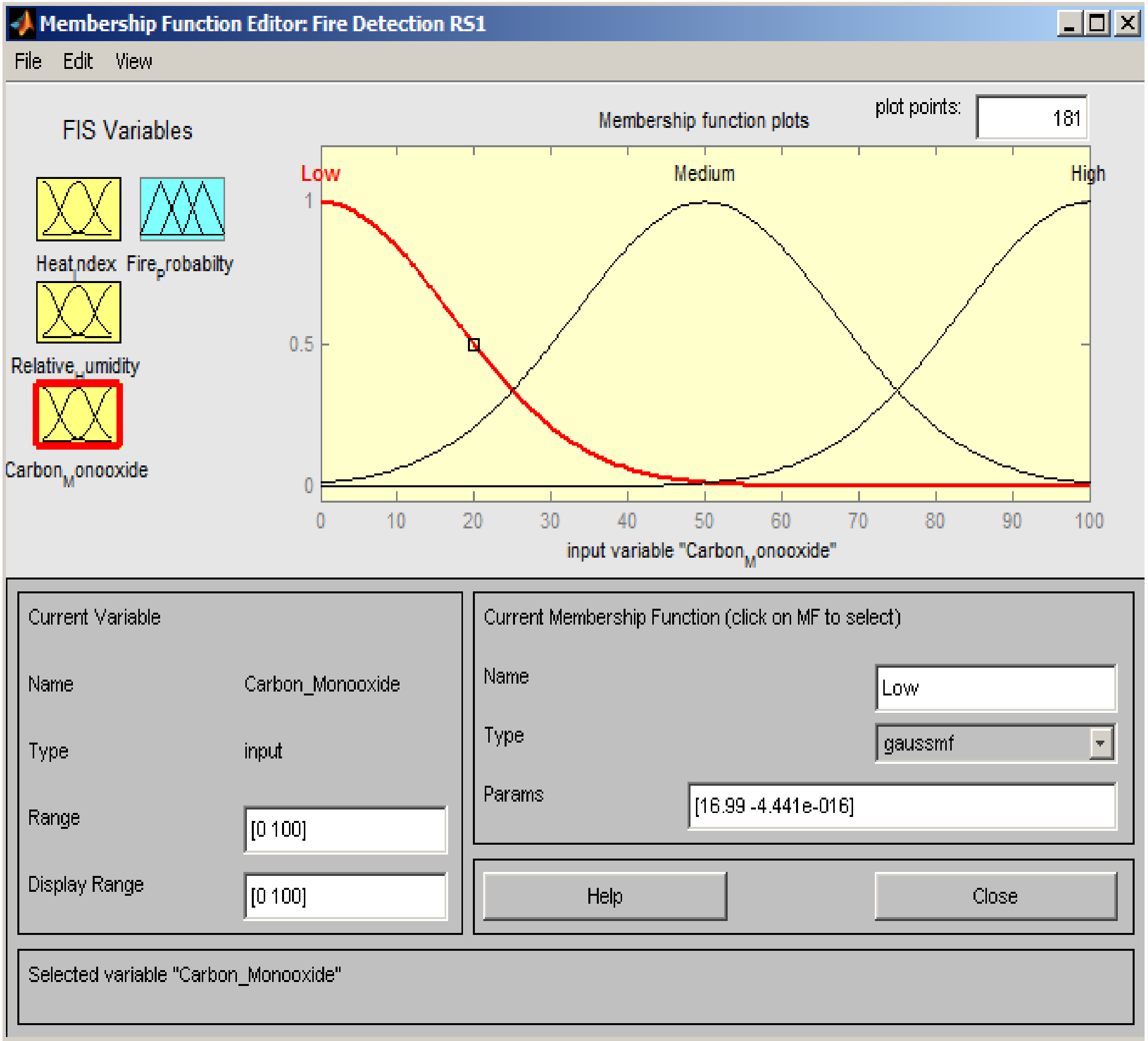


Figure 5.12- Snapshot for input variable Carbon monoxide

Figure 5.12 represent membership function of carbon monoxide. Three mfs are taken named low, medium and high. These mfs are divided over range 0 to 100 parts per million (ppm). Figure 5.13 is showing membership function for output variable that is fire probability divided into over range 0 to 1. The mfs are very low, low, medium, high and very high. Figure 5.14 is snapshot of rule editor. Rules are added here by selecting mfs of input variable and combining them operators. Two types of operator are “AND” and “OR”. We in our experimental evaluation have added 15 rules as listed below. Some rules have been ignored because they were leading towards impossible conditions. Each result contributes to part of final result. Results from each rule are aggregated then final result come out.

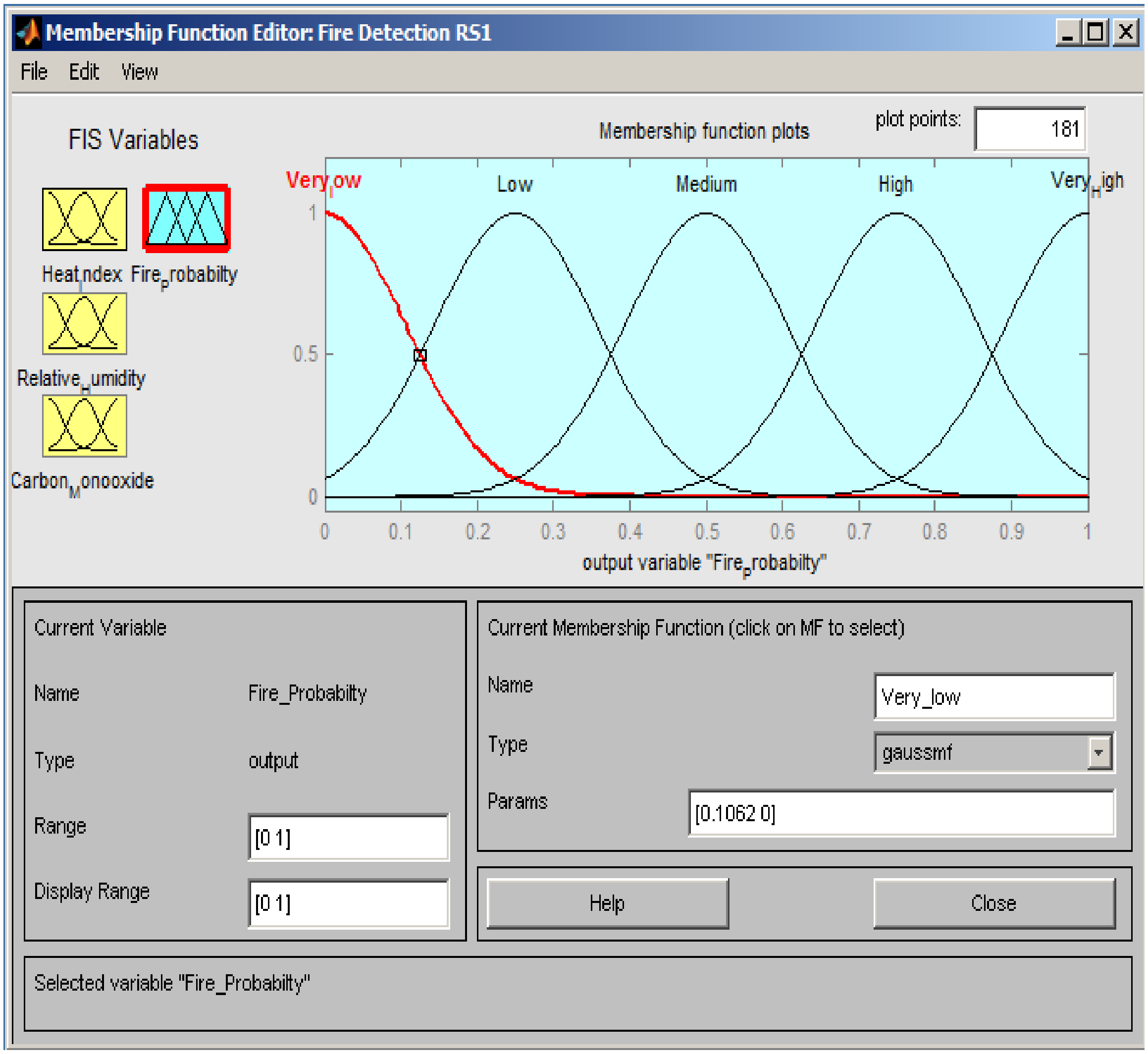


Figure 5.13- Snapshot for membership function of output variable

1. If (Heat_Index is lower) ∧ (Relative_Humidity is Caution) ∧ (Carbon_Monooxide is Low) -> (Fire_Probabilty is Very_low)

2. If (Heat_Index is lower) ∧ (Relative_Humidity is Caution) ∧ (Carbon_Monooxide is Medium) -> (Fire_Probabilty is Low)

3. If (Heat_Index is lower) ∧ (Relative_Humidity is Caution) ∧ (Carbon_Monooxide is High) -> (Fire_Probabilty is Medium). 4. If (Heat_Index is Moderate) ∧ (Relative_Humidity is Caution) ∧ (Carbon_Monooxide is Low) -> (Fire_Probabilty is Very_low).

5. If (Heat_Index is Moderate) ∧ (Relative_Humidity is Extreme_Caution) ∧ (Carbon_Monooxide is Low) -> (Fire_Probabilty is Low)

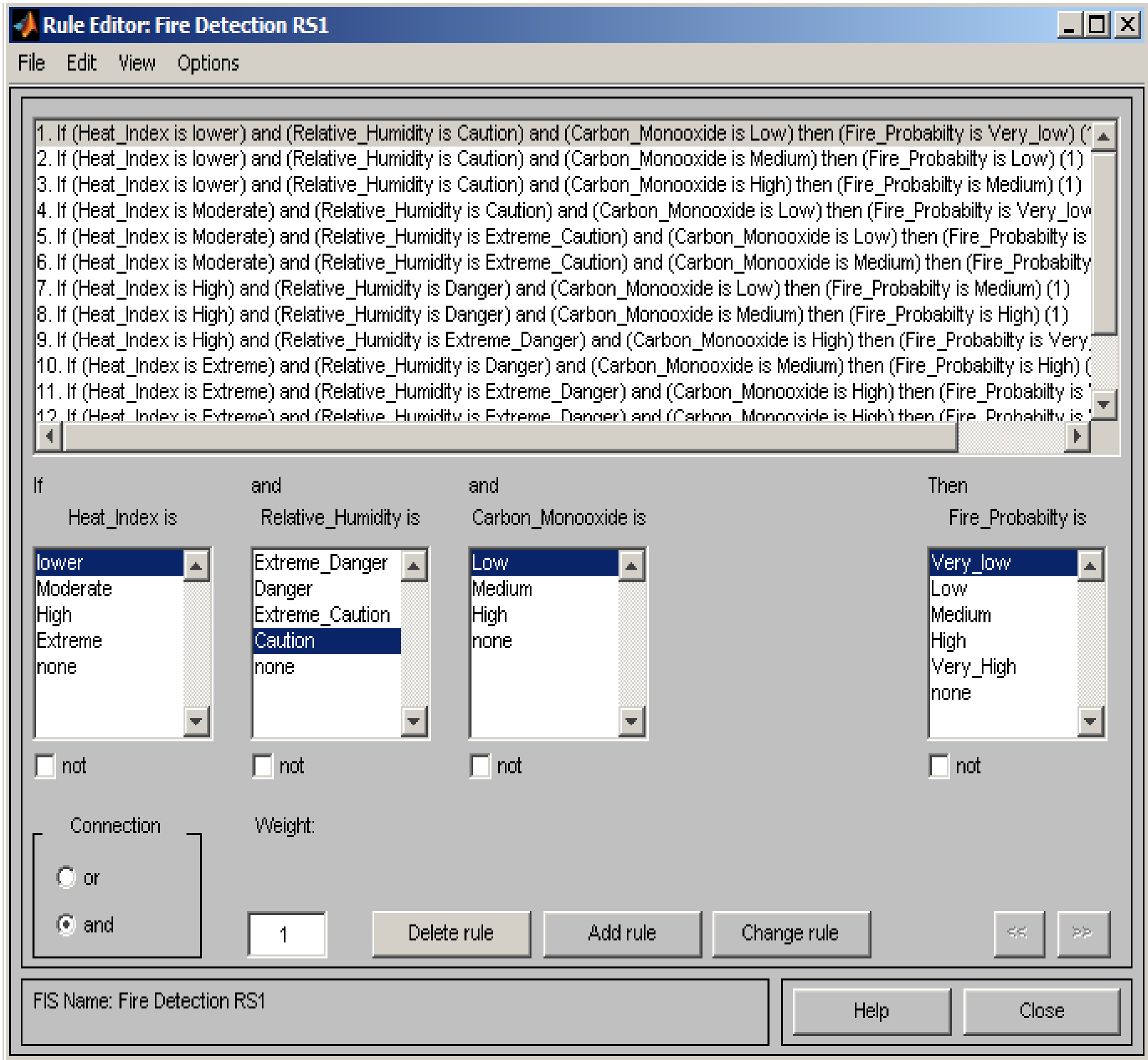


Figure 5.14- Snapshot for rule editor

6. If (Heat_Index is Moderate) ∧ (Relative_Humidity is Extreme_Caution) ∧ (Carbon_Monooxide is Medium) -> (Fire_Probabilty is Medium)

7. If (Heat_Index is High) ∧ (Relative_Humidity is Danger) ∧ (Carbon_Monooxide is Low) -> (Fire_Probabilty is Medium)

8. If (Heat_Index is High) ∧ (Relative_Humidity is Danger) ∧ (Carbon_Monooxide is Medium) -> (Fire_Probabilty is High)

9. If (Heat_Index is High) ∧ (Relative_Humidity is Extreme_Danger) ∧ (Carbon_Monooxide is High) -> (Fire_Probabilty is Very_High)

10. If (Heat_Index is Extreme) ∧ (Relative_Humidity is Danger) ∧ (Carbon_Monooxide is Medium) -> (Fire_Probabilty is High)

11. If (Heat_Index is Extreme) ∧ (Relative_Humidity is Extreme_Danger) ∧ (Carbon_Monooxide is High) -> (Fire_Probabilty is Very_High)

12. If (Heat_Index is Extreme) ∧ (Relative_Humidity is Extreme_Danger) ∧ (Carbon_Monooxide is High) -> (Fire_Probabilty is Very_High)

13. If (Heat_Index is lower) ∧ (Relative_Humidity is Extreme_Caution) ∧ (Carbon_Monooxide is Low) -> (Fire_Probabilty is Low)

14. If (Heat_Index is lower) ∧ (Relative_Humidity is Extreme_Caution) ∧ (Carbon_Monooxide is Medium) -> (Fire_Probabilty is Medium)

15. If (Heat_Index is lower) ∧ (Relative_Humidity is Extreme_Caution) ∧ (Carbon_Monooxide is High) -> (Fire_Probabilty is Medium).

5.6 Results and discussions

Result figure 5.15 is snapshot of rule viewer window. Rule viewer is used to view results. When we vary values of input variables the values we give in form of crisp values in the first step these inputs are Fuzzyfied. Values are mapped to membership functions. Then fuzzy operators are applied using rules. Implication methods are applied. Each implication produces an output. The outputs from all rules are aggregated. The combined output result is then De-fuzzyfied. Because rules results are in fuzzy variable form then De-fuzzification produce crisp values.

FIS take inputs in crisp form and give output also in crisp form. But this mapping is totally based on fuzzy rules in turn also on membership functions. Fuzzy rules are designed on basis of expert's experiences. Here in fire detection FIS we had mapped three input functions to one output function. Three input variables heat index, relative humidity, carbon monoxide are inputted with crisp values. These values are fuzzyfied, mapped to fuzzy output with the help of rules and operators. On basis of probability 3^3 rules can be designed but here some rules were conflicting so we have excluded those rules. Fuzzy rules leads towards fuzzy outputs and fuzzy output is then defuzzified to get crisp value in result form. Here this outputted value is represented as fire probability. Fire probability varied as according to these three values. Figure 5.16 is representing surface viewer.

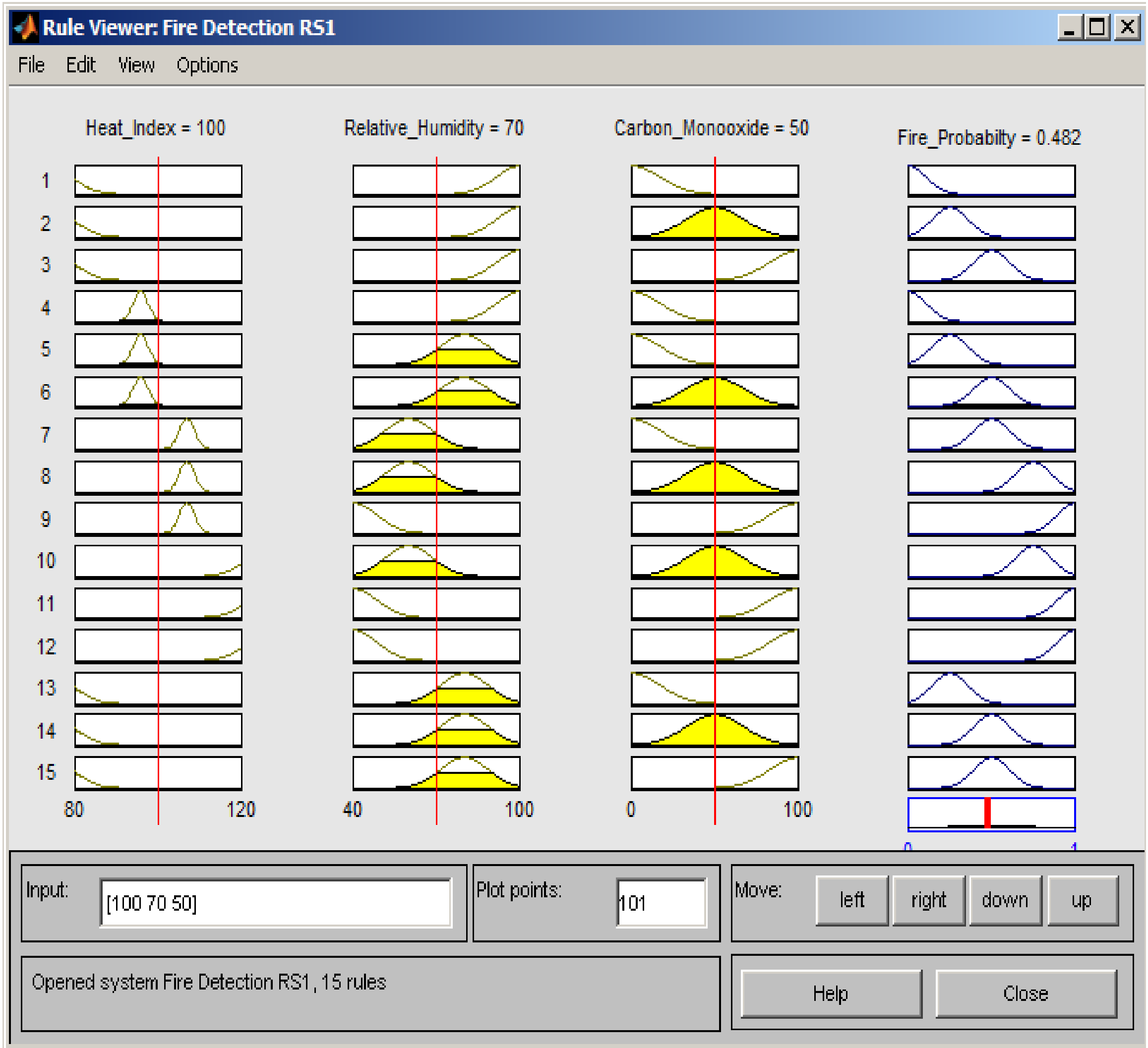


Figure 5.15-Snapshot for Rule Viewer

Result are concluded in table5.2 given below

Table 5.2 Results of Rule Viewer

Input Variables			Output Variable
Heat Index °F	Relative Humidity (%)	Carbon Mono-oxide (ppm)	Fire Probability
85.54	82.3	48.8	0.281
100	70	50	0.482
90.8	59.9	64.5	0.498
110	40	60	0.724
120	50	90	0.85

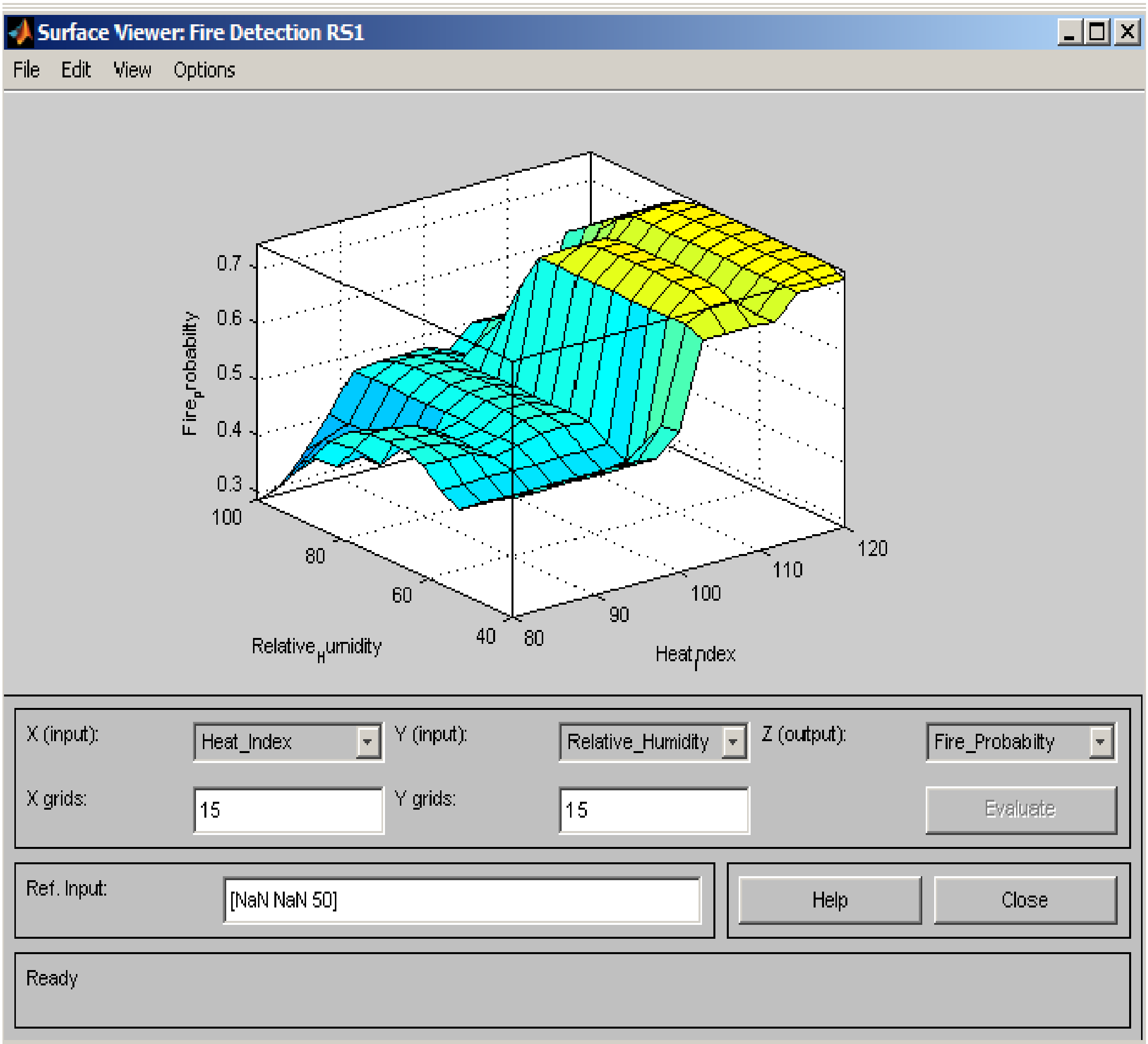


Figure 5.16- Snapshot of surface viewer

5.7 Conclusion

Work has been carried to perform change point detection on event occurrence. Here Change point detection is proposed with MEDC clustering protocol. Change point detection will be done with fuzzy logic approach. The efficiency of this proposed system is checked when we simulated 6 real data sets value on our and 4 temporary sets value for fake fire detection then results was as follows shown in table 5.3. Six real data sets were actual value measured at the time of fire and 4 values were taken randomly. Our aim was to check how much system is reliable and on which probability value it can give maximum efficiency

Table 5.3- Experimental Evaluation on Different Probability values

Pr Values	Total no: of events inputted	No: of event detected	False Detection	Event skipped
0.2	10	10	4	Nil
0.4	10	8	2	Nil
0.5	10	5	Nil	1
0.6	10	4	Nil	3
0.8	10	2	Nil	5

From these results we come up to fact that Pr value 0.5 give 83.3% accuracy; this is the reason why we fixed Pr in our system at 0.5. Hence we conclude that MEDC along with change detection can work well in most cases.

CHAPTER 6 FUTURE DISCUSSIONS

6.1 **Future Work**

- More distribution methods for sensor deployments can be used.

  For sensor deployment we used random function. From literature survey it is found that Gaussian distribution can also save energy and can give benefit to wireless sensor networks. So this sensor distribution method can also be worked with MEDC protocol.

- We can adopt idea of SEECH protocol with our solution.

  MEDC like all other protocols also assumed that cluster heads will be acting as relay nodes. They will forward data which they have received as next hop in multipath communication. SEECH protocol has introduced a different concept that relay nodes will be chased different from cluster heads. So that cluster heads are burdened only with their major role of aggregation and transmission of data which is received from cluster members. Cluster heads should not be burdened for relay services

- More change point detection methods can be used with clustering protocols.

  Many methods exist for information including Fuzzy, Bayesian, Neural networks etc. We have tried to implement fuzzy method for change point detection. Rest methods can also be tried like Adaptive Neuro-Fuzzy may also leads to further perfect results.

- Work can be done with Heterogeneous Nodes also.

  We in network assumptions assumed that nodes are of homogeneous type. Further nodes with different capabilities can be deployed so that some nodes can take privilege then others.

- Chargeable sensors can be taken as further challenge.

  We in network assumptions also assumed that sensors are equipped with non chargeable battery. Work can be carried for chargeable nodes.

Conclusion

This thesis work is aimed to solve two objectives in wireless sensor network; out of which is one is focused on saving energy to prolong lifetime of sensor networks and second objective was to make network more application oriented. Protocols presented in this thesis have improved life time of wireless sensor networks and also contributed a combined solution that will work for clustering and change point detection. By utilizing these protocols, we have also developed change point detections approach which is able to support change point detection in fire application without sacrificing efficiency. These protocols have been validated through the simulation in MATLAB with different scenario

In this thesis we have given three contributory works

- Mutual Exclusive Distributive clustering (MEDC) protocol proposed.
- Mutual Exclusive Hybrid energy Efficient Clustering (MEHEED) protocol proposed.
- Change Point Detection along with MEDC protocol (using Fuzzy Logic)

In first contributory work a new clustering protocol named MEDC. This clustering protocol has shown significant improvement of approximate 33% enhancement in network life time. MEDC is distributive protocol works on message passing system. Idea of MEDC protocol is to select cluster on mutual exclusion algorithm basis. MEDC protocol is having advantage of simple working with least complexity. It selects cluster heads to those sensors which have maximum residual energy within range of communication. On similar network assumption MEDC perform better than HEED the reason of this betterment is; it consider directly residual energy up to last iteration for next iteration's cluster heads selection in contrast of HEED where residual energy is considered for probability calculation. It is also found in worst cases when the sensors with long range of communication are deployed poorly large in number over small region; which may increase in between message communication among the sensors even in these cases MEDC perform equal to HEED. This is disadvantage of message passing system that when the sensor nodes comes closer then large amount of energy is wasted on advertisements instead of transmitting sensory inputs.

Our next contributory work is proposal of MEHEED protocol. MEHEED protocol performs clustering and solves objective of energy saving. It enhance network lifetime of network by reducing energy consumption on communication. This protocol has carried advantage of both protocols one is HEED and other is MEDC. MEHEED protocol posses attribute of fast selection along with simplicity. We have worked and proved its efficiency over previous two protocols. MEHEED perform better in most cases even comparable in worst cases. The reason for be betterment is two fold, one is if selection is based on Chprob then computations are skipped a number of steps are overcooked, on the second hand if selection cannot be taken on probability basis then message passing as like MEDC protocol will select cluster heads in simple way. Message passing will be done and under a range of communication sensor with maximum residual energy will be chased as cluster head. Selected cluster head will transmit cluster head declaration message to sensors under its range of communication. Sensor that comes under range of communication will act as cluster members. Cluster head will aggregate sensory information from cluster members and transmit that data to base station. On maximum exhausted network parameters for all HEED, MEDC, and MEHEED are comparable but when we increase the number of sensors and decrease the range of communication, the communication cost get increased therefore performance get decreased.

Our next contributory work is aimed to make network more application specific more usable; and this is done via introducing change point detection along with clustering protocol. Up to know no clustering algorithm has worked with this change point detection aim. Change point detection is proposed with MEDC clustering protocol. Change point detection will be done with fuzzy logic. When we simulated 6 real data sets value on our proposed system and 4 temporary sets value for fake fire detection then results was as follows. From these results we come up to fact that Pr value 0.5 give 83.3% accuracy; as out of total 6 real events only 1 is missed. This is the reason why we fixed Pr in our system at 0.5. Hence we conclude that MEDC along with change detection can work well in most cases.

List of publications

1. **Urvashi Chugh**, Yashwant Singh, M.V. Ramana Murthy, "Change Point Detection in the midst of Mutual Exclusive Distributive Clustering Protocol" Wulfenia Journal, Klagenfurt, Austria, ISSN 1561-882X, Vol. 22 No.10, pp 311-319, October 2015.
2. Yashwant Singh, **Urvashi Chugh**, M.V. Ramana Murthy "Information Fusion and Change Point Detection in Mutual Exclusive Distributive Clustering" Elsevier Procedia Computer Science; Vol. 65, pp 592-600, 2015.
3. **Urvashi Chugh**, Yashwant Singh, M.V. Ramana Murthy "Comparative Analysis of HEED, MEDC, MEHEED" in proceedings of World Congress on Engineering and Computer Science October, 2015, Vol II, WCECS 2015. ISBN: 978-988-14047-2-5
4. Yashwant Singh, **Urvashi Chugh**, M.V. Ramana Murthy "Clustering, Information Fusion and Event Detection in Wireless Sensor Networks: A Review" in IEEE Conference 9th INDIA Com, March 11-13, 2015, BVICAM New Delhi, pp 288-233, 2015.
5. Yashwant Singh, **Urvashi Chugh**, "MEHEED: Mutual Exclusive Hybrid Energy Efficient Distributed Clustering in Wireless Sensor Networks", Proceedings Engineering Research in Computing, Information, Communication and Applications ERCICA ELSEVIER pp 540-545, August 2014.
6. Yashwant Singh, **Urvashi Chugh**, "Mutual Exclusive Distributive Clustering (MEDC) Protocol for Wireless Sensor Networks", International Journal of Sensors Wireless Communications and Control Bentham Science Press, Vol. 3 No. 2, pp 101-107, 2014.
7. Yashwant Singh, Suman Saha, **Urvashi Chugh** and Chhavi Gupta, "Distributed Event Detection in Wireless Sensor Networks for Forest Fires", IEEE UKSim-AMSS 15th International Conference on Modeling and Simulation, Cambridge University, Emmanuel College, pp 634-640, April 2013.
8. Yashwant Singh, **Urvashi Chugh**, Amit Chugh, "Self Healing Routing in Autonomic Wireless Sensor Networks" Proceedings 6th International Conference on Advance Computing and Communication Technologies, pp 308-312, November 2012.

Under Communication

1. qUrvashi Chugh, Yashwant Singh, M.V. Ramana Murthy, “Performance Analysis of MEDC”, 2016 the 2nd International Conference on Information Technology (ICIT 2016), Melbourne, Australia, March 3-4, 2016. (Accepted) will be published in Journal of Advances in Information Technology **[Indexed in INSPEC, EBSCO, Genamics Journal Seek and World Cat].**
2. Urvashi Chugh, Yashwant Singh, M.V. Ramana Murthy, “Deployment effects on MEDC clustering protocol” communicated Al-Sadiq International conference on Multidisciplinary in IT and Communication Science and Technologies 2016 sponsored by IEEE.