CHAPTER 1

INTRODUCTION

A Wireless Sensor Network is a spatially distributed network consisting of small sensor nodes which communicates themselves using radio signals and are deployed manually or in a random fashion to sense, monitor and understand the required area. The sensor nodes forms a network which sense the desired data like environmental conditions and physical conditions such as sound, temperature, pressure, movement etc and the sensed information is transmitted to the central entity called Base Station (BS). The BS is also called sink.

As the advancement in the technology, the wireless sensor network is used in various fields of applications. The size of the sensor nodes is small and nowadays there is no fixed topology for the deployment of these sensor nodes in the environment which makes utilization of these nodes more efficient [1].These features increased the applicability of wireless sensor network in an un-attended harsh environment such as military applications, security surveillance, weather monitoring, vehicular movement, health and industry applications [2,3].There are limitations of the sensor nodes which includes limited power battery , low bandwidth, less memory space , low processing capabilities.

A lot of study and research work has been done in area of WSN. The major domain of research is to increase the lifetime of the network. To increase the lifetime the number of transmissions between the nodes should be minimized. There are many protocols used to increase the energy efficiency of the WSN's which is based on the architecture of the nodes which transmits information between themselves. These include routing protocols, clustering protocol, data aggregation protocols, data fusion protocols, data compression protocols, etc.

The WSN comprises of data acquisition network and data distribution network. These two networks are controlled and managed by a central unit.

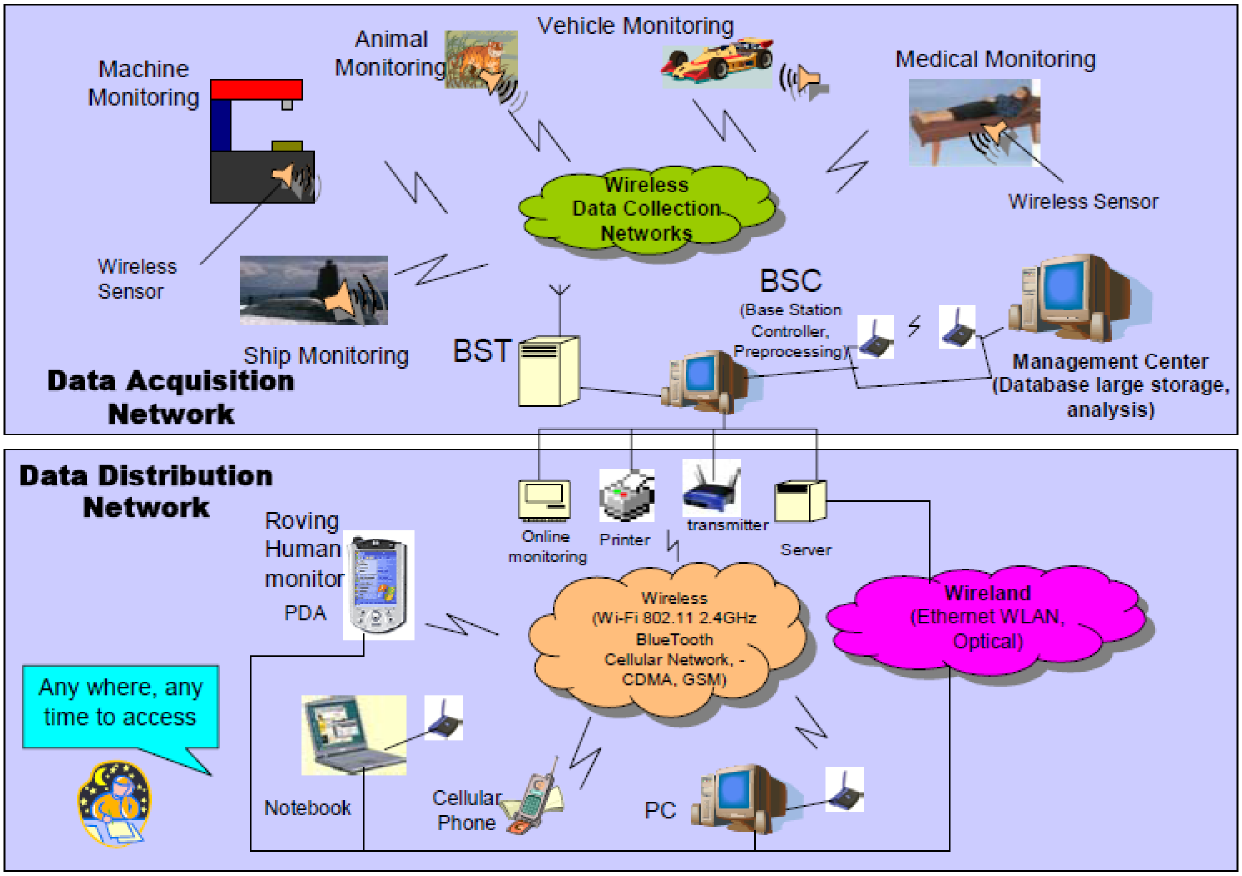

Figure 1.1: Wireless Sensor Network

Figure 1.1 represents the basic wireless sensor network. The sensor nodes gathered the data from the environment by monitoring (animal monitoring, vehicle monitoring, medical monitoring, ship monitoring). The collected data is sent to the base station at which processing is done. The base station can store this data in a large database for further analysis. The processed data is sent to the clients who request for the desired information. These clients can be any cellular phone, notebook, printer, any transmitter or server can relay information. And this information can be accessed at any location and at any time by the clients.

1.1. Basic Architecture of Sensor Node

Sensor nodes are low powered highly integrated digital electronics and micro-mechanical-electro system. The circuitry senses the conditions of the environment and transformed into an electrical signal.

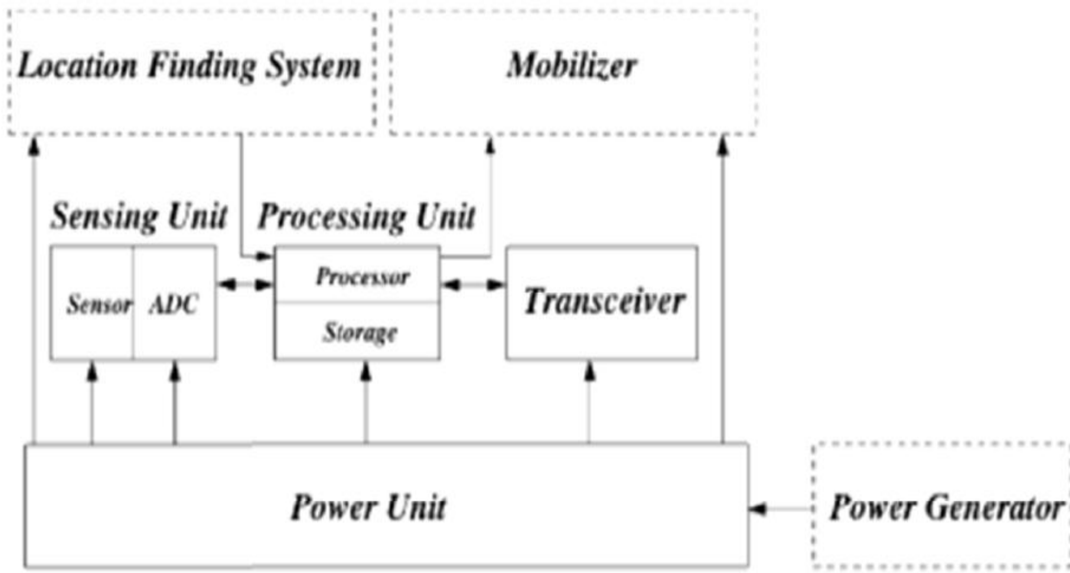

Figure 1.2: Sensor Node Components

Figure 1.2 represents the basic structure of the sensor node. There are four basic components of a sensor node: sensing unit, processing unit, transceiver unit and power unit [2]. A mobilizer and a location finding system is also present. The sensing unit has two main components: sensor and ADC (Analog to Digital Convertors). The analog signals produced by the sensors are converted to the digital signal by ADC. These digital signals act as an input to the processing unit. The processing unit consists of processor and storage. The transceiver unit is responsible for the connection of the node to the network. The power unit is an essential unit and it supplies power to all other components to accomplish their required task. These nodes are non rechargeable but there are some sources which can assist this unit like solar cells. There is an optional location finding system unit which helps in localization of the nodes and this information is used in many routing applications. Thus major research is done to make use of power unit efficiently and prolong the lifetime of the network.

1.2. Sensor Network Protocol Stack

The sensor network protocol stack has some planes in additional to the layers of TCP/IP to deal with the issues of the sensor nodes.

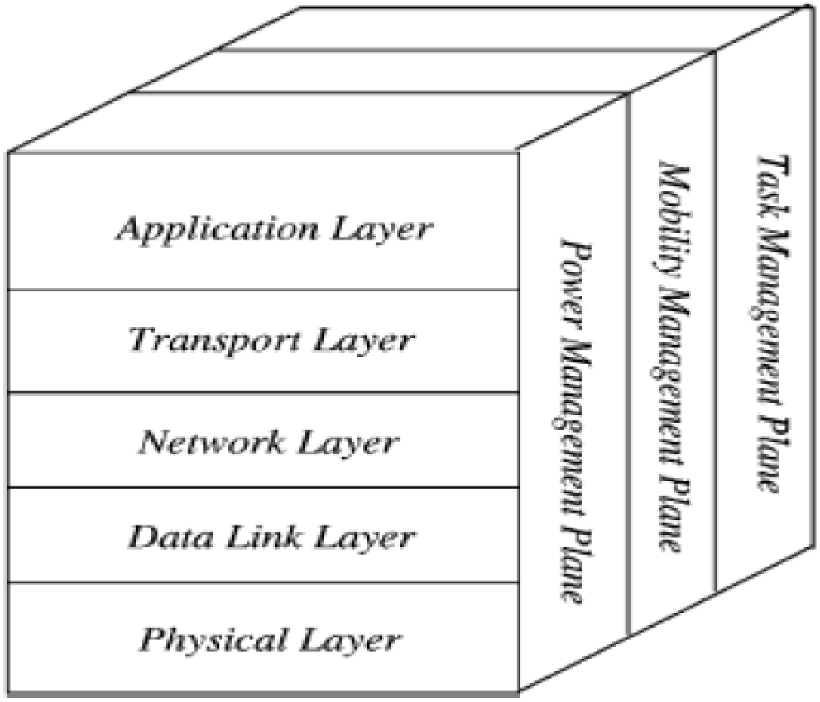

Figure 1.3: Sensor Network Protocol Stack

The sensor network includes the traditional layers which has the same functionality as in TCP/IP. The Application layer is responsible for the development of application software which is based on some sensing tasks. The transport layer manages the flow of the data. The data is provided to the network layer by transport layer and network layer performs the task of routing. The noisy data and mobility of the nodes is handled by the data link layer. Data link is aware of the power management. While transmitting or broadcasting the data it helps in reducing the collision with the neighboring nodes. The transmission, modulation and receiving the signal is handled by the physical layer. The additional planes deals with mobility, power management and task management which includes the needs of the sensor network. The Power Management Plane (PMP) helps in deciding whether to switch on the sensor node or switch off the sensor node (when not in use). It also decides whether to transmit information or not. All the nodes do not sense and transmit the information all the time so to decide when and whom to sense all these decisions are made by Task Management Plane (TMP). Mobility Management Plane (MMP) deals with knowing the path to base station and also informs the newly updated location of the neighbors. Without these additional planes the objective of the WSN is not satisfied. These planes provide a way to conserve the energy of the sensor nodes. These planes take care of the changes in the mobility, power and task of the sensor nodes.

1.3. Characteristics of WSN

There is a large number of sensor nodes involved and these can be placed static or dynamically in the region. Some characteristics are given below:

- Energy aware usage: As energy is the serious issue in the WSN and to increase the lifetime of the network. The nodes are made to sleep for particular period of time or when they do not have any new information to send to the base station.
- Self configuration: As the topology of the network tends to change every time due to the dynamic nature of the network. The nodes adapt themselves if any new node is added or any previous node is deleted or failed in the network.
- Single hop communication: In single hop communication the node sends it data to base station directly. But this leads to transmission of redundant data and it is also an inefficient way of sending data. If every node sends its data this leads to more energy consumption.
- Multi hop communication: In multi hop communication the node sends data to intermediate nodes and these intermediate nodes send data to the base station. Multi hop communication is used in the cases when distance between the node and the base station is large or base station is beyond the coverage or transmission range of the sensor node. This process of packet forwarding is very efficient and energy conserving.
- Load balancing: As the nodes have limited power, any node can be dead anytime so another node can take charge of that node to continue the work.

1.4. Motivation

Clustering in wireless sensor network is a very important task because it affects the functionality of the whole sensor network. The wide use of wireless sensor network in real world now in form of different applications like border surveillance where every aspect of the network is very critical because even a short glitch in network functionality can result in tremendous after effects, it is very important to work on the problems of WSN and make it more usable and reliable.

There are different evolutionary and population based optimization techniques applied for clustering like Particle Swarm Optimization (PSO) [22] and Genetic Algorithm (GA) [20], Ant Colony Optimization (ACO) [21]. In recent time, a new swarm based optimization algorithm called Firefly algorithm is proposed by Xin-She Yang[26]. Firefly algorithm performs better as compared to other well known population based techniques like PSO and GA.

1.5. Objective

The objective of the work is explained as:

- To find the optimal number of cluster heads and their associated cluster members.
- To form clusters with two cluster heads per cluster.
- To increase the number of alive nodes after rounds of communication.
- Compare our protocol which takes into consideration all the above characteristics with the conventional clustering protocols in terms of energy efficiency.

1.6. Thesis organization

We start this dissertation with introduction in chapter 1. A detailed description of background is presented in chapter 2 which includes applications, issues and characteristics of WSNs. Chapter 3 describes the nature inspired clustering protocols in detail which are related to our research problem. Chapter 4 gives a brief about the network model we have used. Chapter 4 also explains in detail about our proposed algorithm Firefly-DCH. We evaluate the performance of the proposed algorithm with other clustering protocols in chapter 5. Conclusion and future work is explained in chapter 6.

CHAPTER 2

LITERATURE REVIEW

2.1. Classification of WSNs

Sensor nodes are deployed according to their applications in various fields. In some cases the nodes are randomly distributed as the area may be isolated. But in other cases the nodes are statically deployed according to some predefined method. Due to the different usage of the WSN the structure of the network differs from application to application. Some of the classification of WSN [4 5 6] is discussed as:

2.1.1. Aggregating and non aggregating network

On the basis of data aggregation capability of nodes the network is categorized as aggregating and non aggregating networks. In non aggregating network the node directly sends the data to the BS. This technique is suitable when size of the network is small. But when there is large number of nodes aggregating network is used. In aggregating network the node makes use of the intermediate nodes to pass the data to the BS. Sometimes the BS is not in the coverage of the node so the node transmits the data first to the nearby node and that node then transmits it further. The aggregating network reduces the traffic in the network which ultimately improves the performance of the network. In non aggregating network the redundant data is transmitted in the network. So when the density of the network is large aggregating network is preferable.

2.1.2. Single hop and Multi hop network

Single hop and multi hop network is differentiated on the basis of hop count. In single hop network the data is directly transmitted to the BS in single hop. But in multi hop network the data is transmitted with the help of the intermediate nodes. For example there are five nodes A, B, C, D, E and a BS. Figure 2.1 shows the basic structure of single and multi hop network. In single hop network all the nodes send their data directly to BS but in multi hop network node A and C send its sensed data to node B and node E send its sensed data to node D. Now node B and node D send this received information to BS.

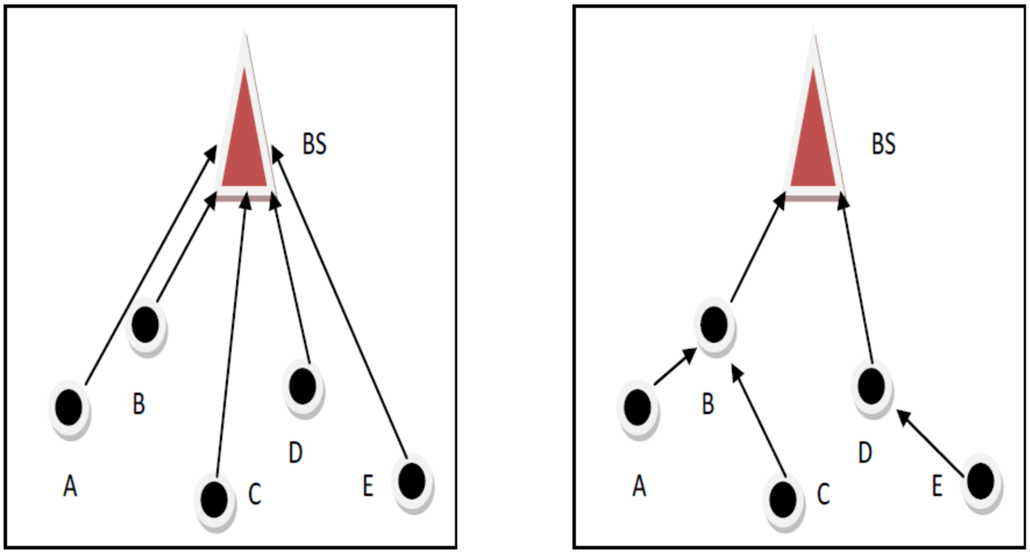

Figure 2.1: Single hop and multi hop network

2.1.3. Flat and hierarchical network

In flat architecture all the nodes in the network have same capabilities and each node plays the same role of performing sensing task. All the nodes in the network act as peers. In hierarchical architecture all the sensor nodes within the network is organized into clusters. Within the cluster all the members send their information to head and the head then send data to the sink or BS. Figure 2.2 and 2.3 shows the architecture of flat and hierarchical network respectively.

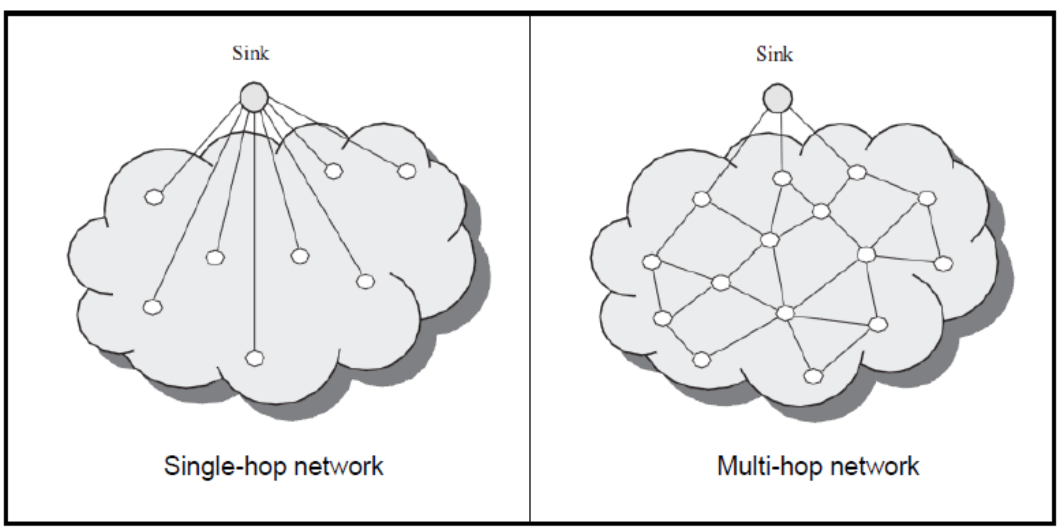

Figure 2.2: Flat architecture networks

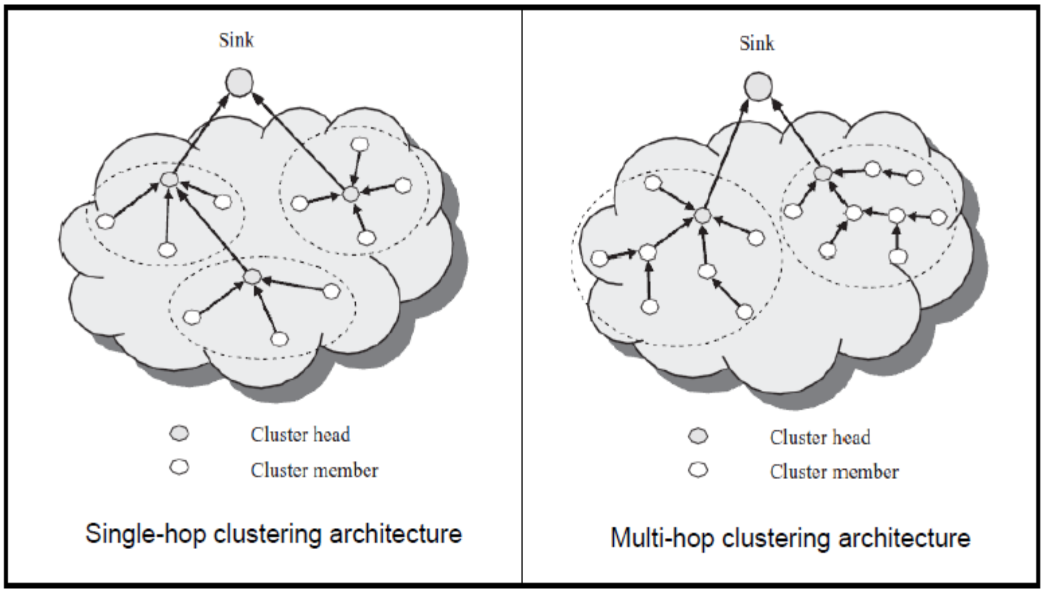

Figure 2.3: Hierarchical architecture networks

2.1.4. Deterministic and dynamic network

This classification of network is based on the method of node deployment. In areas such as forest, under water, or any area in which human intervention is not possible dynamic deployment of nodes is suitable. In deterministic network the topology is preplanned and the nodes are deployed on the fixed positions.

2.1.5. Reactive and Proactive network

There are some tasks which require continuous transmitting of sensed data but on the other hand some tasks need information at a particular time. In proactive network the sensed information is transmitted to the intermediate node or BS periodically. At constant interval of time the information is sent to the Base Station. But in event driven applications the use of proactive network leads to wastage of resources so reactive network is used. In reactive network the sensors continuously sense the environment and send the information only when some specified event occurs like change in temperature, pressure etc.

2.1.6. Self configurable and non self configurable network

Nowadays the use of WSN is increasing in dynamic environment so the self configuration tendency of the network is required. Because the topology is not predefined so to handle changes in the environment self configurable network is used. There is continuous addition and deletion of nodes in the WSN so to manage all these changes self configurable network is preferred. The non self configurable network is basically used when the topology is definite and there is less probability of changing the topology and also when the size of the network is small.

2.2. System Evaluation Metrics

Now the key metrics that are used in evaluation of the WSN are explored. According to the usage and the objectives of the network deployment is kept in mind [7].

- **Lifetime**: The wireless sensor network is used in environmental monitoring and security applications and these are left unattended for months and years. So the lifetime of the network becomes important in this scenario. And the primary limiting factor for the lifetime is supply of energy, so each node must manage its energy supply. Some power supplies from outside can be provided like in form of solar cells or piezoelectric generators [8]. But both these demands low power consumption nodes.
- **Coverage**: This is also a primary evaluation metric. The physical area for the WSN is very large and it is required to deploy the nodes in this large network. It also possible that a node is very far away from BS (sink) or the BS is not in the transmission range of the node. So to enable the full coverage of the network multi hop scheme is used.
- **Cost and ease of deployment**: One of the major advantages of using WSN is the ease of the deployment of nodes in critical environment. The self configuration nature of the WSN made successful deployment of nodes. The relocation of the nodes over a large physical area is a continuous process. And the objects may be placed in between and can interfere with communication between the two nodes. So self maintenance is required.
- **Response time**: As in alarm application scenario, alarm must be signaled immediately when an intrusion is detected. So the message should be transmitted with high priority.
- **Scalability**: The WSN consists of hundreds of thousands of nodes so the network should be scalable so that network can respond to certain events efficiently.
- **Fault tolerance**: There are some conditions which lead to failure of nodes like decay of power, interference by an obstacle, physical damage to the node. So despite of being in such situations the network should perform as specified and the working of the network should not be affected. Fault tolerance can be defined as the ability of the network to work as expected even after failure of node.

- **Operating Environment:** The WSN can be formed in dangerous environments like beneath the sea, in the chemical reactor, on borders, forest etc. The network can be attached to any person or animal or installed in high speed vehicles.
- **Data delivery models:** Based on when to send the sensed data to BS there are some models defined. These models depend on application and the models are event driven, continuous, query driven, hybrid. In event driven model the sensed data is transmitted to BS when a particular event occurs. In continuous model the data is sent by the nodes continuously to the BS. In query based model the data is sent when the BS send some query to the nodes. In hybrid model some of the previously defined approaches of models are combined.
- **Data fusion:** There are many sensor nodes that sense the environment and collect the same data. So to minimize the transmission of redundant data over the network data aggregation is used. In this technique the data from various nodes are combined with the help of some functions and then this data is sent to the intermediate node or BS. This requires some computation but it saves lots of energy of the network by reducing the number of transmissions.
- **Overhead and data latency:** Multi hop transmission and data fusion adds latency in the network. There are some overheads which are attached to complex routing protocols that are not suitable for the energy constraint network.
- **Quality of service**: quality of service is dependent on the type of the application. It may be reliability, energy efficiency, awareness of location, etc. Depending on different applications different protocols are used like in military application the data needs to be sent continuously and in addition to this security is also an important aspect.

2.3. Application areas of WSN

With the advancement in the field of WSN, it is used in many fields of applications. The different types of sensors used in sensor network are seismic, low sampling rate, magnetic, thermal, visual, infra red, acoustic, radar which monitor different conditions like temperature, humidity, vehicular movement, pressure, noise level, soil makeup, mechanical stress etc. Figure 2.4 shows overview of the applications of WSN. Some of the applications are discussed as:

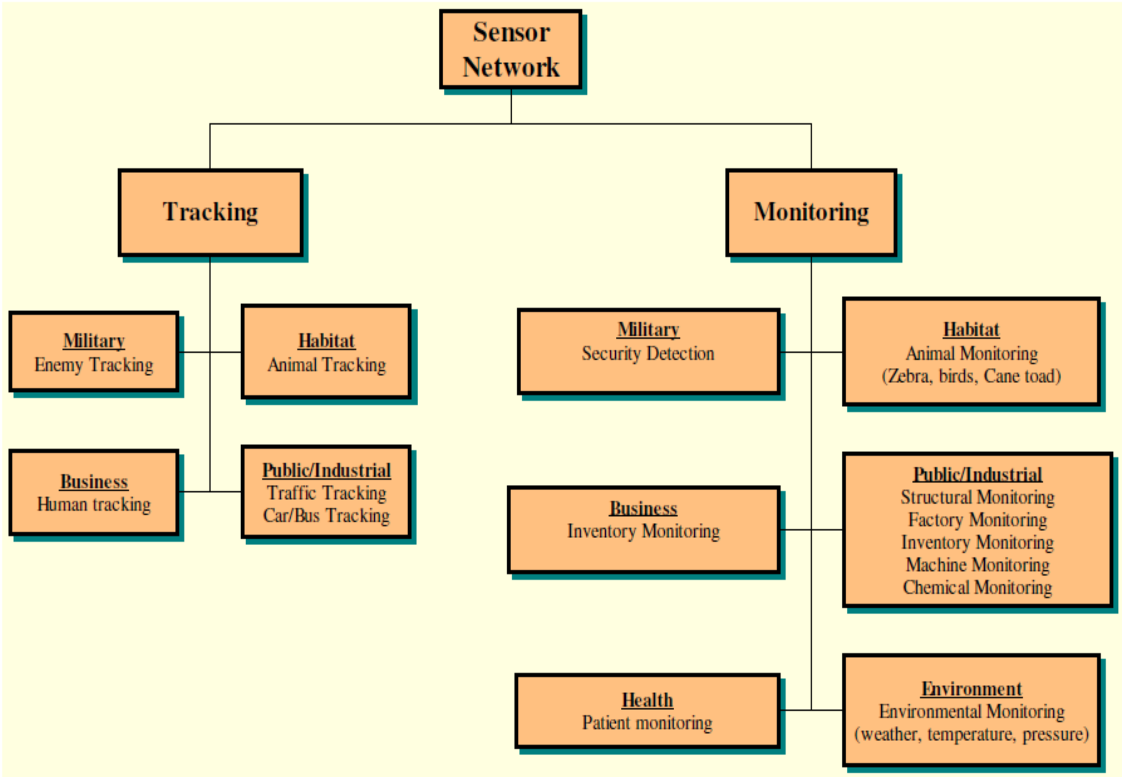

Figure 2.4: Sensor Network applications

2.3.1. Health applications

WSN has been widely used in health industry which covers diagnostics, providing interfaces to the disabled, patient monitoring, drug administration, monitoring the movement of internal process of insects, tracking patients and doctors inside the hospital [9].

The physiological data collected by the sensor network is stored for a longer period of time that can be used for further consultations by the experts.

2.3.2. Military applications

WSN is mostly used military applications like communication, computing, intelligence, surveillance etc [10]. The various features of WSN like self configurable, ease of deployment, fault tolerance all these are required in military applications. There is a need to set up a quick network and the nodes may deploy as fast as possible. To accommodate the changes in the network the self configurable property is very helpful in these high tension areas. The military operations involve monitoring the arms and ammunition, surveillance of the battlefield, guidance for the target, detecting the chemical and biological attack. The activities of the opposing forces can be closely watched. The assessment of the data gathered after the attack can be estimated quickly and after analyzing the sensed data nodes can be deployed in the target area. WSN can be used for monitoring the troops of the friendly forces to keep a close watch on them. WSNs can also be used in managing the information related to the availability and conditions of the equipments used in the battlefield before and after the battles.

2.3.3. Environmental applications

WSN is widely used in monitoring the environmental conditions which are related to natural calamities and disaster. Other application involves tracking the movement of animals; bio complexity and pollution study, monitoring the condition of the crops, monitoring soil and atmosphere, detecting forest fire and flood [11].

For flood detection ALEART system is deployed in US. WSN is helpful in monitoring the levels of pesticides in drinking water, level of soil erosion and air pollution. The data collected by the weather, water level, rainfall sensors are sent to the BS and accordingly the BS take actions. In forest fire detection the sensors that are randomly deployed can find the exact location of the origin of the fire and send this critical data to the BS before the fire becomes uncontrollable. On the basis of the received information further steps can be taken which can save human losses.

2.3.4. Commercial applications

This application includes monitoring product quality constructing smart office, automatic manufacturing, interactive toys, and automation in factory process, machine transportation, and robot control. Some of the factory applications are factory instrumentation, control of actuators, instrumentation of semiconductors processing chambers, rotating machinery, wind tunnels.

The temperature control of the building or offices is maintained by the WSN as there is difference in temperature at different places. The central unit maintains the temperature in these kinds of offices. The various kinds of theft can be monitored and detected and helps to report this activity to the end user. The inventory in the warehouse can be managed by the sensors attached to it. The location of the item can be easily found and if some insertion is performed then the user need to attach the sensor node to the inventories and count of the items can be easily maintained.

2.4. Clustering in Wireless Sensor Network

There is a large number of energy constraint nodes involved in forming the WSN and their batteries are also non rechargeable. For high scalability and increased lifetime some special energy aware routing and data gathering protocols are used. To satisfy the scalability issue the sensor nodes are grouped into clusters. And this approach of combining nodes together is widely used by the research community. The organization of the clusters is also supported by the hierarchical routing and data gathering protocols. So that data fusion and aggregation is possible that can save energy of the whole network. Figure 2.5 shows the basic cluster formed in the WSN.

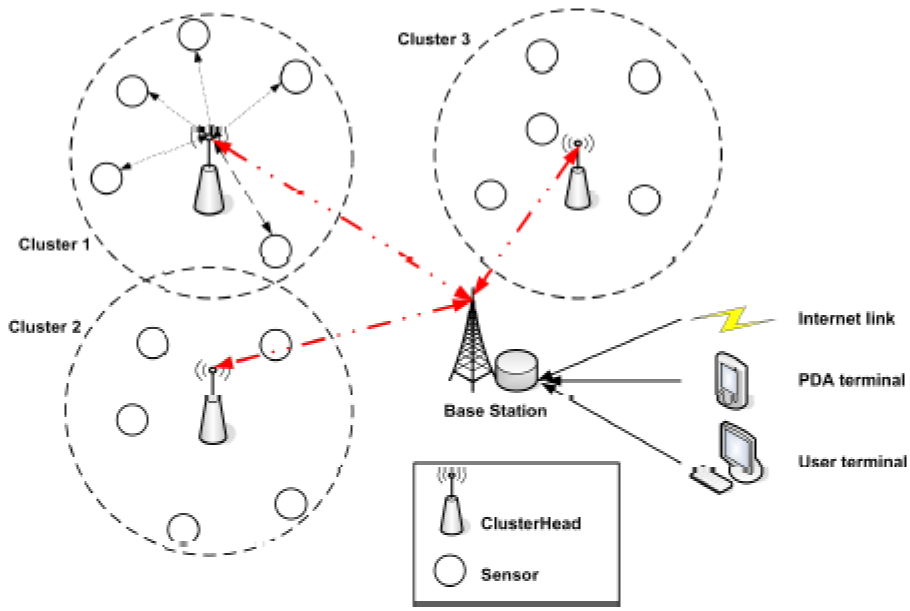

Figure 2.5: Clustering in WSN

The hierarchical network consists of various clusters and each cluster has special node called CH (Cluster Head) and remaining nodes in the cluster acts as member nodes. The CH has some special duties to perform within the cluster. Then the network is divided into two levels one level is called higher level which consists of CH nodes and another level is called lower level consists of member nodes. The member nodes transmit their sensed data to their corresponding cluster head.

Then CH performs some computation task on the data received from nodes and then transmits to the BS. The CH acts as intermediate between BS and the nodes. The CH perform transmission over a large distance so its energy degrades faster than the member nodes so to balance the energy cluster head is re elected in the network.

2.4.1. Objectives and design challenges

The hierarchical clustering provides scalability, lifetime and energy efficiency to the network. Hierarchical routing contributes in low energy utilization within the cluster by performing data aggregation so that less number of messages is transmitted through the network. While in non clustered approach each node sends its sensed data to the BS directly which can causes traffic within the network which ultimately injects latency in communication. Additionally, the non clustered approach is also not suitable for the large area of network and is not scalable for these large areas as the nodes have limited range of transmission. The clustering also supports load balancing and efficient resource utilization. There are some other advantages also. The localization of the nodes and the routing decisions are easily done and the routing table size also reduced. The topology within the cluster can be implemented on the basis of the nodes available or as per the application. The nodes within the cluster do not have to worry about the structure of the network outside the cluster. As there is limited number of nodes within a cluster so the maintenance of the nodes is easy. The CH can schedule the nodes within the cluster and can switch the nodes from on state to off state and thus eventually increasing the lifetime of the network. The transmission of the redundant messages and collision rate within the network is greatly reduced.

The basic challenges faced are limited energy, limited capability, and network lifetime. The additional considerations in design process of clustering are formation of cluster, synchronization, data aggregation, dependency on application, security. In formation of cluster it is important to select the appropriate cluster head and adequate number of clusters. The number of messages exchanged within the cluster must be minimum and it should be independent of the size of the network. Within the cluster synchronization must be maintained between the exchanges of messages. There must be any scheme specified like TDMA for the nodes to send messages and when to go in sleep mode. Data aggregation deals with the fusion of data and also minimizes the energy in the network. The clustering and routing protocols are designed in such a way so that it can adapt according to the application requirement. Security is an aspect which must be met in networks like military applications.

2.4.2. Parameters of clustering

The different parameters are discussed as:

- Number of clusters: The number of clusters formed affects the other protocols like routing, data gathering etc. The selection of cluster head and number of cluster are sometimes predefined according to the application. In other conditions the number of cluster can be dynamically determined.
- Intra-cluster communication: The communication between the nodes and its CH can be direct or indirect. Sometimes multi hop intra-cluster communication is done in cases where the CH nodes are limited.
- Mobility: When all the nodes are stationary then cluster formation is easy and cluster and cluster heads become fixed. But if the nodes are moving, then the node will changes the cluster and its membership will becomes dynamic.
- Types of nodes: In homogeneous network, all the nodes are assumed to have same capability. In heterogeneous network, the CH nodes are assumed to have more computation and communication capabilities than the other nodes.
- Methodology: In centralized approach there is some central authority which decides the formation of clusters but nowadays all the nodes participate in the decision of selecting CH and this approach is distributed in nature.
- Selection of cluster head: The cluster head can be selected using any deterministic algorithm or may be selected using any probabilistic algorithm or may be in completely random manner.
- Algorithm complexity: The algorithm chosen for the forming the cluster must converge as fast as possible and should have fewer complexes in nature.
- Levels: There may be hierarchy of clustering which results in multiple levels of clusters. Multiple levels are used in very large networks and where energy efficiency is a critical aspect.
- Overlapping: Some protocols supports overlapping to some extent but other does not support overlapping of clusters at all.

2.4.3. Classification of Clustering Protocols

There is some common classification for the algorithms that are used for clustering [12]; clustering algorithm for homogeneous and heterogeneous network, centralized or distributed algorithms, static and dynamic clustering.

The first classification is based on the characteristics and functionality of the nodes in the network. In heterogeneous sensor network two kinds of nodes are present; one has greater capabilities and is predetermined to be CH, other nodes just acts as members and sense the data from the environment. In homogeneous network all the nodes have same capabilities and each node is capable of becoming the CH. The CH role can be rotated among different nodes in the homogeneous network to achieve better load balancing.

The other classification is based on the method used to form the clusters. In centralized approach the CH is selected by the coordinator or BS. But this is only suitable for small networks, for large networks distributed approach is required, in which the selection of CH is based on all the nodes. In static formation of clusters the CH is predefined but in dynamic formation the CH is re elected due to the topology changes or to accommodate the changes in the network. Re election of CH can be sometimes performed on specific occurrence of events. Dynamic clustering is better than static clustering in terms of network lifetime and energy efficiency.

Another category of clustering algorithm is based on the method of formation of cluster and parameters used for selection of CH.

2.4.3.1. Probabilistic Clustering algorithm

In probabilistic clustering algorithm, probability is assigned to each sensor node and is used to select the appropriate CH [13]. Some other features such as residual energy, transmission cost, distance between nodes is also considered. This protocol uses random or hybrid methodology. This clustering method is fast and has less number of messages exchanged within the network.

2.4.3.1.1. Low Energy Adaptive Clustering Hierarchy (LEACH)

LEACH is the most popular protocol used for clustering in WSN [14 15]. It is the first protocol that is based on the needs of the WSN. It is distributed, self configurable, adaptive, randomized algorithm. This protocol assumes that the sensor nodes are stationary and randomly deployed within the network. It is a hierarchical protocol which aims to improve the lifetime of the network by distributing the consumption of energy between all the nodes. The clusters are formed on the basis of received signal strength (RSS). All the further processing and decisions like routing are taken within the cluster.

The clusters are formed by the autonomous decision of the nodes without the interference of any centralized control. Each node has equal chances of becoming the CH so that the energy spent in every round is balanced. Initially the node decides to be a CH with probability "p" and broadcasts its decision. After the election each CH broadcast an advertisement message to the other non CH nodes and these non CH nodes now selects the cluster based on the distance between itself and CH. There is rotation of in the role of the CH which is performed periodically to balance the load. Each chooses a random number "T" between 0 and 1. A node is selected for the current round if the number chosen is less than the threshold.

$$T(i) = \{ \quad \frac{p}{1-p(rmod(1\backslash p))} \qquad if\ i\ \mathcal{E}\ G \tag{2.1}$$

If i does not belong to G then it is updated as 0.

p is the probability of each node,

r is the current round number,

G is the set of nodes that are been selected for CHs in last 1/p rounds.

In each round the clusters are formed dynamically and the time to perform the next round is also chosen randomly.

There are some disadvantages which include wrong selection of CH as the node with low energy may be selected. It is possible that the selected CHs may not be evenly distributed within the network. There is a possibility that the selected CH is not in the transmission range of some nodes. The sensor node is selected as CH which is far from the BS or not in the transmission range of BS. LEACH allows only one hop communication which is not suitable for the large wireless sensor networks.

2.4.3.1.2. Hybrid Energy Efficient Distributed Clustering (HEED)

HEED is distributed and hierarchical clustering protocol [16] and allows multi hop communication between CHs and BS. Unlike LEACH, HEED does not select CH randomly. The basic parameters that are involved in selecting CH is residual energy of the node and intra-cluster communication cost. This algorithm evenly distributes the CH within the network and works in three steps. In the first step number of CH formed is selected and it is used to limit the initial number of CHs. The second step involves fixed number of iterations; each sensor node performs these iterations until the CH is selected. If no CH is heard then the sensor node chooses itself as CH and sends the control message to other nodes informing this change. Each CH doubles its probability and goes to next iteration and the iteration stops when the probability becomes 1. In the final step either the sensor chooses its CH or itself becomes the cluster head. In large sized network HEED performs better than LEACH.

2.4.3.2. Non probabilistic clustering algorithm

The non probabilistic clustering algorithms are more deterministic in terms of selection of CH and forming the clusters as compared to probabilistic algorithms. The main focus of these algorithms is the information received from the neighboring nodes and the distance between the neighboring nodes. The messages exchange between nodes is more so results in more time complexity. In addition to proximity of the nodes some criteria like residual energy, transmission range, and mobility are also considered.

2.4.3.2.1. Node Proximity and Graph based Clustering Protocol

This is distributed multi hop hierarchical clustering algorithm which forms multi level clusters [17]. In this algorithm any sensor node can initiate the cluster formation process and it is performed in two steps: “Tree Discovery” and “Cluster Formation”. The initiator node performs the distributed Breadth First Search (BFS) for the formation of the tree. The node u broadcast its shortest hop count to the root, r. The node v, chooses node u as parent when its distance is shortest to root r via node u. The broadcast signal contains parent ID, root ID, sub tree size. Every node changes the sub tree size when its children sub tree size changes. When the size of the sub tree of a node increases in size and becomes more than the predefined size parameter k, the cluster formation starts. The size of the cluster is 2k.

2.4.3.2.2. Weight Based Clustering Protocols

This algorithm considers residual energy, combined weight, transmission range in addition to the node proximity for the formation of clusters [18]. This is a weight based protocol which selects CH based on “combined weight” W, for each node which includes certain parameters. These parameters involve node degree, residual energy, transmission power, mobility. This information is broadcasted by each node and the node chooses the smallest weighted node as its CH. If topology of the network changes then again weights are calculated this process of cluster formation is repeated. This provides better load balancing but achieving synchronization is an important task.

CHAPTER 3

NATURE INSPIRED PROTOCOL

Wireless Sensor Technology is evolving rapidly in every field of application. Now the research community is focusing on the nature inspired algorithms to be used in WSN. There are various issues related to the WSN which are efficiently handled by the biologically inspired protocols. In large scale networks clusters helps in data gathering and routing. And using bio-inspired algorithms to form clusters is an efficient way of clustering in WSN. The nature inspired is meta-heuristic, multi objective, stochastic algorithm. These algorithms are suitable for the optimization problems [19]. Optimization is a technique which deals with finding the best solution for a given problem. When the search space is very large then the previous algorithms do not perform as expected. But in this scenario the nature inspired algorithms outperforms the other protocols. These algorithms are suitable for random search and can easily find global optimal solution. The main concept of these algorithms lies in its iterative property. These algorithms work best when there is mutual cooperation required between the nodes. As the networks keeps on changing its structure these algorithms are suitable and helps to adapt according to the environment always.

The traditional algorithms stuck in the local optimal solution and have high time complexity when search space becomes large. These traditional algorithms are not capable of performing specified task when number of nodes increases. In clustering the nature inspired algorithms are widely used and also resolve the problems faced by the traditional approaches. The accuracy, speed and convergence are the specific parameters that are related to these algorithms. These algorithms are scalable in nature and have less chances of getting stuck in local optimal solutions. As these are adaptable to the environment so the self configurable characteristic is also inherited.

The WSN or the new architectures are including the bio-inspired algorithms because of the following reasons:

- The bio-inspired algorithms are easily adaptable according to the environment.
- These algorithms are highly fault tolerant and have resistance to internal or external failures.

- These algorithms allow complex situations to be implemented in limited set of basic rules.
- These systems are able to learn and evolve as new conditions occur..
- These algorithms perform efficiently under limited resources.
- These algorithms are capable of self configuration in distributed environment for efficient collaboration.

3.1. Clustering using Genetic Algorithm (GA)

GA is widely used in WSN for energy optimization. In clustering, the GA is used to decide the number of clusters and the number of members within each cluster. The algorithm is used in selecting the CH and also helps in scheduling the transmission of data [20]. The algorithm has two phases cluster formation step and data transmission step. The output of the GA is the selected CH in the network. In this algorithm the distance between the sensors nodes are calculated and in addition to this distance to the sink are also determined. The main step is the evaluation of the fitness function of the problem which is dependent on various parameters. The sensor nodes are represented as bits of a chromosome. The head is represented by 1s and members are represented by 0s. There are various parameters that are associated in calculating the fitness of a chromosome like number of nodes, energy of a node, and distance between the nodes. The solution contains chromosomes and these chromosomes form a population. Then the best chromosomes decide the next population of the solutions. The initial population contains a large number of CHs which are randomly chosen. The next generation is formed on the basis of the fitness of the previous generation. If there is n number of sensor nodes in a WSN then it is represented by a chromosome of m bits. The objective of the fitness function is to reduce the energy consumption within the network and to increase the lifetime of the network. The fitness parameters involve distance of the node from the sink and cluster head, energy used by a node in transmission, the number of total transmissions performed by the node and standard deviation of the cluster distance. The fitness function of a chromosome involves all these fitness parameters. Initially the weights assigned to these parameters are random. And in further iterations the best chromosome is evaluated and weights are adjusted accordingly. The fitness parameter of a chromosome is updated after each generation and is evaluated for improvement in number of clusters, cluster size, energy consumption of the network.

3.2. T-ANT Protocol for clustering

The T-ANT protocol is formed on the basis of the Ant Colony Optimization (ACO). The main objective is to find the optimal number of clusters and to select the evenly distributed CHs in the network. In this each sensor node is associated to only one CH [21]. There are some assumptions related to design:

- The protocol is distributed in nature and every node can take its decision.
- There is constant time used in determining the CH.
- The complexity should be less in terms of time and processing.
- The CHs are evenly distributed.

The algorithm consists of two phases namely cluster setup phase and steady state phase. It is a dynamic protocol which has a certain number of rounds and each round includes the above two phases. In cluster setup phase the clusters are formed by selecting the appropriate CH and the neighboring node are associated to the selected CHs. In steady state phase the data is transmitted to the CH and the process of data gathering is done. This gathered data is then transmitted to the sink at regular interval of time. There is an agent used in the protocol that is called ants and these ants play a major role in finding the solution.

A fixed number of ants are released by the sink during initialization. These ants randomly choose any neighbor and then the ant performs a random walk through the network. The ants release a special substance called pheromone which is used for the communication purpose between them. But at this stage the ants do not release pheromone to avoid attraction between them. The next ant is released by the sink after a randomly chosen interval of time. The ant chooses the next node randomly and then moves towards that node. If two ants coincide at the same node then the latter ant has to change the location and find the other node.

A timer is started in the cluster setup phase to select the CH and to form the cluster. When the timer expires the node checks whether it has an ant associated with it. If there is an ant then that node will be selected to become a CH. A fixed number of CHs are formed in this phase as there are fixed number of ants are used. And this algorithm do not uses looping of any function that contains number of nodes so the selection of CH is terminate in constant time. Now the selected CH will advertise the message to the other nodes and based on the distance between the nodes and the cluster head, the node chooses the CH with minimum distance. This algorithm is particularly used when the convergence should be fast and there is need for load balancing in the network.

3.3. Clustering using Particle swarm Optimization (PSO)

PSO is technique based on the collective behavior of natural species as herd of animals or flock/swarm of birds [22]. The PSO is a population based strategy which optimizes the objective function. The method involves collection of search points called particles and calculates fitness of each particle. Each particle has an associated velocity; this velocity helps in moving towards the better solution if the cost of the fitness function is optimized. The particles have corresponding fitness value for all the positions they visit. In each iteration the local best fitness value of the particle is maintained and this information helps in reaching towards global best position. During each generation, each particle uses the information about its previous best individual position and global best position to maximize the chances of moving towards the better solution space which results in better fitness. If better solution is found as compared to previous one then the solution is updated according to the following equations where t represents time:

$$\mathrm{vid}(t) = w * \mathrm{vid}(t-1) + c1\emptyset1\big(\mathrm{pid} - \mathrm{xid}(t-1)\big) + c2\emptyset2(\mathrm{pgd} - \mathrm{xid}(t-1)) \quad (3.1)$$

$$\mathrm{xid}(t) = \mathrm{xid}(t-1) + \mathrm{vid}(t) \quad (3.2)$$

vid: the particle velocity, xid: the particle position, $\emptyset1, \emptyset2$: random numbers;

w: inertia weight c1, c2: learning factors; pid, gid: particle′s best and global positions

PSO is widely used in clustering in WSN for energy efficiency. PSO reduces the cost of locating optimal position for the head nodes in a cluster. The clustering is performed in a centralized manner using PSO [23]. In the centralized approach the BS implements the clustering and in distributed approach the sensor nodes performs clustering. All the sensor nodes send their information about energy level or the distances with different nodes to the sink. The aim is to optimize the objective function and objective function includes the residual energy, intra-cluster distance. The algorithm assumes number of particles in the sensor area and iterates in search of the best location for the cluster head. The velocity parameter changes as the position of the particle changes. In this the sensor node is stationary but only the position of the particle changes. The node which is found nearest to the particles is associated with that particle and its residual energy is taken into consideration. The BS maintains the local best position of each sensor along with other characteristics. The algorithm operates in rounds and these rounds include setup phase and steady state phase. The setup phase involves PSO and selects the cluster head. In setup phase each sensor node in the network broadcasts HELLO packet which contains its id. The sensor node that receives the packets updates its neighbor table and includes that received id into its table. This is done for the purpose of neighbor discovery. After the neighbor discovery is done by all the sensor nodes in the network the nodes now send the information to the BS. The information contains the node id, residual energy and neighbor table data. Based on the information received by the BS, the BS calculates the average energy of all the nodes. The nodes with the energy above the average energy level will be considered for the selection of CHs. Then the BS will run the PSO algorithm to find the best CHs. The particle size in the PSO is equal to the upper bound on the number of CH candidates. After the selection of the CH based on the PSO algorithm, the clusters are formed. In the cluster formation the non CH nodes are assigned to CH nodes according to the distances between them, the nearest CH is chosen by the non CH sensor node. In state phase the CH becomes the controller of the other nodes and coordinates the data transmission within the cluster. The messages are then sent to the CH and to avoid collision the CH sets up a TDMA schedule for its members. To further reduce the energy consumption of the sensor nodes, CH allows the sensor nodes to turn on their radio devices when they are transmitting and to turn off when they are not transmitting information to the BS. When the CH collected sensed data from the sensors, it processes the data and sends it to the BS. There are some types of PSO clustering discussed as follows:

3.3.1. MST-PSO: Minimum Spanning Tree PSO

The clustering algorithm of the weighted graph in the WSN is sometimes based on the minimum spanning tree PSO [24]. The optimal route between the sensor node and the CH is searched from the entire optimal tree; this search is based on the energy consumption. The CHs are selected on the basis of the energy of the sensor nodes and Euclidean distance with its neighbor node in the optimal tree. It is proposed that once the network topology is decided the network lifetime becomes settled. There are two techniques suggested for improving network lifetime are firstly reducing the start up energy consumption and secondly optimize the network topology.

3.3.2. Distributed PSO

The PSO algorithm deals with minimizing the energy consumption maintaining connectivity between the nodes. Some methods are proposed [25] for network topology that deals with hidden nodes and asymmetric links. To increase the lifetime of network and to minimize the asymmetric links and hidden nodes some subsets of node are made to transmit more power. Author proposed algorithm to generate topologies with less number of hidden nodes and asymmetric links and it is proved that his topology deliver more data and last longer.

CHAPTER 4

PROPOSED WORK

This chapter gives the understanding of the proposed work and along with the details of the algorithms. In this chapter we discuss firefly algorithm and how it used in clustering in WSN. Firefly algorithm is recent algorithm which under swarm intelligence developed by Xin-She Yang [26] in late 2007 and 2008 at Cambridge University. It is stochastic; meta-heuristic that can be used to deal with optimization problems. The heuristic here means to find the solution with trial and error method. This algorithm is inspired by the flashing behavior of fireflies in nature. It is also efficient in solving multimodal problems. It solves the problem by searching for the solution in a randomization manner. These algorithms are able to solve NP hard problems also. The algorithm searches the new solutions in the search space and considers the best solution out of it. To avoid the solution to be stuck in local optima, the algorithm has to maintain balance between exploration and exploitation. Exploration has the ability to jump out of the local optima and search the space globally and exploitation is using the local knowledge which is gathered so far in the search process and the next search is concentrated on the local neighborhood where the optimality may be close. Exploitation tends to increase the speed of the convergence and exploration tends to decrease the rate of convergence of the algorithm. The main characteristic is the flashing light of the fireflies and the fireflies use this light for two purposes one is to attract the mating partners and second is to warn predators. The light intensity I decrease as the distance r increases; this obeys the inverse square law. This is the basic phenomena which motivated the use of firefly algorithm in various fields and also in the field of WSN. Mutual coupling occurs between the fireflies if two fireflies are in the vicinity of each other. This mechanism of their communication using flashes and their synchronization is being studies and used in various techniques of WSN. The multi modality solving tendency of this algorithm makes it suitable for solving objective function which have different functions or dependent on various other functions.

4.1. Structure of the firefly algorithm

There exists a biochemical process called bioluminescence which produces light flashes in fireflies. This bioluminescent process takes place in an organ called lanterns. The algorithm is based on the formula of light intensity I that is decreases with increase in the distance. Due to the absorption of light, it becomes weaker and weaker as distance increases. Firefly Algorithm has three rules which are define below:

- All the fireflies are unisex which implies that one firefly will be attracted to other fireflies.
- The attractiveness is proportional to the light intensity and both of these decreases with as the distance increases. So the less bright firefly will move towards the brighter firefly. The firefly will move in random direction if there is no brighter one than itself.
- The light intensity of the firefly is dependent on the objective function.

The light intensity of the firefly represents the solution which is dependent on the fitness value of the objective function. The variation in the light intensity with distance r is defined as:

$$I(r) = I_0 \exp(-yr^2) \tag{4.1}$$

I_0 denotes the light intensity of the source and y is the light absorption coefficient. Similarly the attractiveness β is proportional to light intensity I(r), hence the attractiveness is defined by the equation:

$$\beta = \beta_0 \exp(-yr^2) \tag{4.2}$$

β_o refers to the attractiveness at r=0.

The light intensity is the absolute measure of the light actually emitted by the firefly and the attractiveness is the relative measure of the light that is seen by the other fireflies.

The firefly algorithm is used iteratively to solve the optimization problem, there is n number of fireflies and x_i denotes the solution for the firefly I which has its fitness value. The distance between two fireflies' x_i and x_j is calculated as the Euclidean distance.

$$rij = \sqrt{(\sum (xik - xkj)^2)} \tag{4.3}$$

As k varies from 1 to d and d denotes the dimensionality of the problem. The movement of the firefly i towards the brighter firefly j is defined as:

$$xi = xi + \beta 0 \exp(-\gamma (rij)^2)(xj - xi) + \alpha \epsilon \tag{4.4}$$

In the above equation, first term is the position of the firefly, second term is due to attraction and third term induces randomization, where α is a randomization parameter and $\mathcal{E}_i$ is a vector of random numbers drawn from uniform or Gaussian distribution. If the value of β_0 becomes zero then it becomes a random walk. The parameter y deals with the convergence speed. The firefly algorithm basically has three controlling parameters: the randomization parameter α, the attractiveness parameter β and the light absorption coefficient y.

4.1.1 Firefly algorithm and complexity

The basic firefly algorithm is summarized as:

```
Objective function f(x), x=(x1, x2, x3 ...xd) T
Initialize a population of fireflies xi (i=1, 2, 3.........n)
Define light absorption coefficient y
while (t<MaxGen)
    for (i=1 to n)                   // for all fireflies
       for  (j= 1 to i)
             Light intensity Ii at xi is calculated by f (xi)
             If (Ii<Ij)
                 Move firefly i towards firefly j                    // acc to (4.4)
             else
               Move firefly i randomly
            End if
         Attractiveness varies with distance r                       // acc to (4.2)
         Evaluate new solutions and update light intensity
         end for j
     end for i
 rank the fireflies and find the current best
end while
```

The firefly algorithm has two inner loops for covering the entire population and one outer loop for specified number of iterations. So the complexity at the extreme case is O (n^2t). If the n is small and t is large then algorithm complexity is linear in terms of t and is relatively inexpensive. For the optimization problems, all the complexity and the computationally extensive part is calculation of objective function.

4.2. System Model

In this network model and energy model are discussed which are used in the proposed algorithm. The network model lists the details of the capabilities of the sensor nodes and the working environment of these nodes. The energy model discusses the usage of the energy of the sensor nodes when these nodes perform different tasks.

4.2.1. Network Model

The network contains n number of sensor nodes and one BS, the dimension of the network area is D X D and the BS is present in the network area only. The nodes are deployed randomly in this network and after deployment the nodes are stationary. The sensor nodes continuously sense the data from the network and send the data to the BS. The BS analyzes the data and provides it to the end user according the requirement of the application. All the nodes are homogeneous in sensing tasks and all the nodes have limited source of energy and their batteries cannot be recharged. This is the network model which is under consideration for the proposed protocol.

4.2.2. Energy Model

The first order radio energy dissipation model [27] is used as the energy model for the sensor nodes. The power amplifier, the transmitter, the receiver are the three main components. The energy is dissipated in these components to run the electronic circuitry. The transmitter uses energy to run the transmitter circuitry, the amplifier transmits data packets and the receiver uses energy to run the receiver circuitry while receiving the packets. The figure 4.1 shows the basic first order radio energy model.

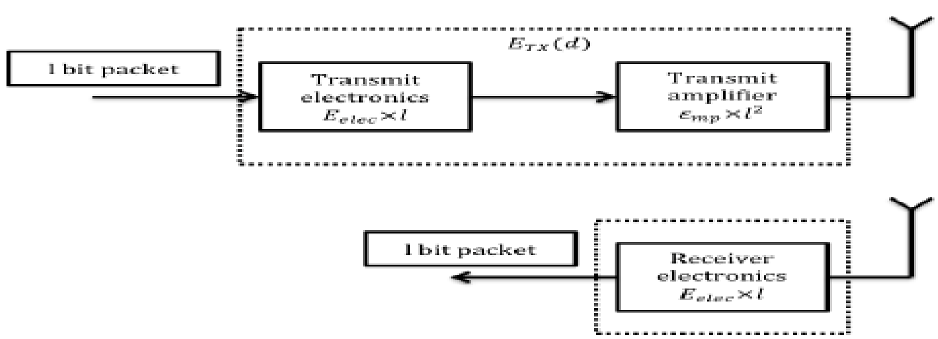

Figure 4.1: First order radio energy model

The energy dissipation is dependent on the distance of transmission. The energy dissipation is inversely proportional to square of distance if distance is small between sending and receiving nodes, otherwise energy dissipation is inversely proportional to d^4. The energy consumed in transmitting k bits of information at a distance d with acceptable Signal to Noise ratio (SNR) is:

$$ETx(k, d) = \begin{cases} Eelect * K + \mathcal{E}fs * K * d^2 & \text{if } d < d0 \\ Eelect * K + \mathcal{E}amp * K * d^4 & \text{otherwise} \end{cases} \quad (4.5)$$

Where ETx is the energy dissipation in source node transmitter, Eelect is the energy consumption per bit to run the transmitter or receiver circuit, Ɛamp and Ɛfs are the amplifier parameters. The d_0 is calculated as:

$$d0 = \sqrt{\left(\frac{\mathcal{E}fs}{\mathcal{E}mp}\right)} \quad (4.6)$$

To receive k bit of information the energy dissipates by the node is:

$$ERx(K) = Eelect * K \quad (4.7)$$

4.3. Clustering Using Firefly Algorithm

The firefly algorithm is proved to be efficient in performing clustering in WSN [28]. Performing clustering using firefly algorithm is a centralized approach in which the sensor nodes send data to the BS. The BS will analyze the received data and select CH by applying firefly algorithm and then the sensor nodes will form clusters.

As we know this is centralized scheme so all the decisions are taken by BS and BS will run firefly algorithm to select the K number of best CHs that minimize the objective function. The objective function is the cost function which is the combination of two individual functions and this whole cost function is to be minimized in this algorithm.

$$cost = \beta * f1 + (1 - \beta) * f2 \quad (4.8)$$

$$f1 = \max\left(\frac{\sum d(ni, CHp,k)}{|Cp,k|}\right) \quad (4.9)$$

$$f2 = \frac{\sum E(ni)}{\sum E(CHp.k)} \quad (4.10)$$

Where k varies from 1 to K. The f_1 function is the maximum average Euclidean distance between sensor nodes and their corresponding CHs. $|C_{p,k}|$ is the total number of nodes that belongs to the cluster C_k of particle p. The f_2 function is the ratio of the total initial energy of all the sensor nodes to the total number of candidate cluster heads in the current round. The β is a constant that is fixed by the user to make sure how much weight should be given to the sub objectives of the cost function. If equal weight is given to both the sub functions then the value of β is set to 0.5. The objective function is always dependent on the problem definition or the characteristics of the problem domain. As it can be clearly seen that residual energy should be considered while selecting the cluster head and the distance between the CH and cluster member should also be considered as the cluster member sends its sensed data to the associated CH.

4.2.1. Cluster Setup

Assume the WSN has N number of nodes and K number of predefined CH and then clustering can be done in WSN by following the below algorithm step by step.

1.) Initialize s number of particles so that it contains K number of randomly selected CH chosen from the total number of cluster head candidates.

2.) Cost function of each particle is calculated as:

   i.)For each node n_i $i=1, 2, \ldots\ldots\ldots\ldots N$

   Calculate distance d (n_i, $CH_{p,k}$) between node n_i and all the $CH_{p,k}$ candidates.

   Select the cluster head $CH_{p,k}$ for the node n_i and assign this n_i to the $CH_{p,k}$ which has minimum distance as $d(n_i, CH_{p,k}) = \min\{d(n_i, CH_{p,k})\}$ for $k= 1, 2, \ldots\ldots K$.

   ii.)Calculate the fitness value using the cost function // acc to (4.8) to (4.10)

3.) Rank the fireflies according to the fitness value and find the current best firefly.

4.) Update the position of the particle // acc to (4.4)

5.) Limit the change in the particle's position value.

6.) Map the new updated position to the nearest (x, y) coordinate.

7.) Repeat steps 2 to 6 until the stopping condition is met (i.e. the maximum number of iteration is reached).

Now the BS has the set of optimal CHs and their corresponding cluster members. The BS transmits this information to all the sensor nodes in the network.

4.3.2. Problem definition

We have focused on the area of clustering in WSN. There is one CH selected per cluster and the CH is overloaded with the messages received from the cluster members. As there is receiving, processing and transmitting operation of information is involved at each CH within the cluster. The energy of the CH depletes very fast or it can be situation in which CH die suddenly due to external or internal issues. So in order to maintain the proper functioning of the network, the CH is selected again by performing another clustering round. And this re-clustering is time consuming and a lot of energy is wasted.

4.3.3. Solution: Double Cluster Head

One of the solutions to minimize the energy and time spent in re-clustering is proposed as instead of selecting one CH in a cluster we have selected two CHs namely primary CH and secondary CH. The primary cluster head and secondary cluster head will perform all the specified tasks which are done by the normal CH. The sensor node within the cluster sends its data to the CH which is nearer to it. This double cluster head strategy will make the network more fault tolerant and the network continues to work in situations when the selected CHs within the cluster breaks down. This will increase the working capacity of the network.

The previous firefly algorithm outputs only one CH per cluster but this modified firefly algorithm will produce two CHs, one is primary and other is secondary.

4.3.3.4. Modified Firefly Algorithm for Clustering

1. Using LEACH algorithm performs clustering and selects the initial cluster heads.
2. Based on this information, the initial cluster head runs the firefly algorithm to select the primary cluster head and secondary cluster head.

```
Generate S particles to randomly contain K number of Cluster heads.
Map the randomly generated positions with the nearest (x, y) coordinates.
While (t<MaxGen)
Map the positions with the nearest (x, y) coordinates.
Evaluate the cost function of each particle.                  // acc to (4.8) to (4.10)
    For (i= 1 to n)                                           // for all fireflies
      For (j=1 to n)

           If (I_i <I_j)
                Update the particle's position                // acc to (4.4)
                 Limit the change in the particle's position value
            End if
```

End for

End for

Rank the fireflies according to the cost function and select the current best.

End while

Select the global best as primary CH and, the global best of previous iterations as secondary CH.

Now the BS has the optimal set of primary CHs and secondary CHs and the BS will send this information to all the sensor nodes in the network.

This Firefly-DCH seems to be a favorable optimization tool phenomenon due to the effect of the attractiveness function, which is unique to the firefly behavior. Firefly not only includes the self improving process with the current space, but it also includes the improvement among its own space from the previous stages.

We have simulated this double CH technique and observed that it increases the lifetime of the network.

CHAPTER 5

SIMULATION RESULTS AND ANALYSIS

Simulation is considered as efficient and flexible tool to evaluate the performance of the algorithm working under vivid environmental conditions. The performance of the algorithm (Firefly-DCH) is compared with other conventional protocols in terms of number of nodes alive after certain number of rounds.

5.1. Simulation Setup

Matlab is the tool we used for simulation and performance evaluation of the Firefly-DCH. Our aim with the simulation is to compare LEACH and standard firefly algorithm with the modified firefly algorithm in respect to number of nodes alive after predefined number of rounds.

There are 100 nodes present which are uniformly distributed in the network. The dimension of the network field is 500m * 500m. All the 100 nodes are deployed randomly in the 500m * 500m field dimension network. The BS is located at the centre of the network i.e. at location (250, 250). First order radio model is used as the energy model as described in the previous sections. There are 80% nodes that have 2J energy and 20% of nodes have 5J energy. The number of iterations performed by firefly algorithm for selecting CHs is 150. The number of rounds used for the simulation is 1000 rounds and length of the data packet is assumed to be 4000 bits for non cluster nodes and 8000 bits for cluster heads. The values of the parameters used in the simulation are explained in the table 5.1

Parameters	**Values**
E_{elec}	50nJ/bit
ε_{fs}	10pJ/bit/m^2
ε_{amp}	0.0013pJ/bit/m^4
N	100
R	1000

Table 5.1: Parameter Setting

5.2. Performance Evaluation

In this section we simulate our proposed algorithm into Matlab environment and compare with LEACH and standard firefly algorithm. We have also implemented the concerned algorithm in the same simulation environment for the purpose of comparison.

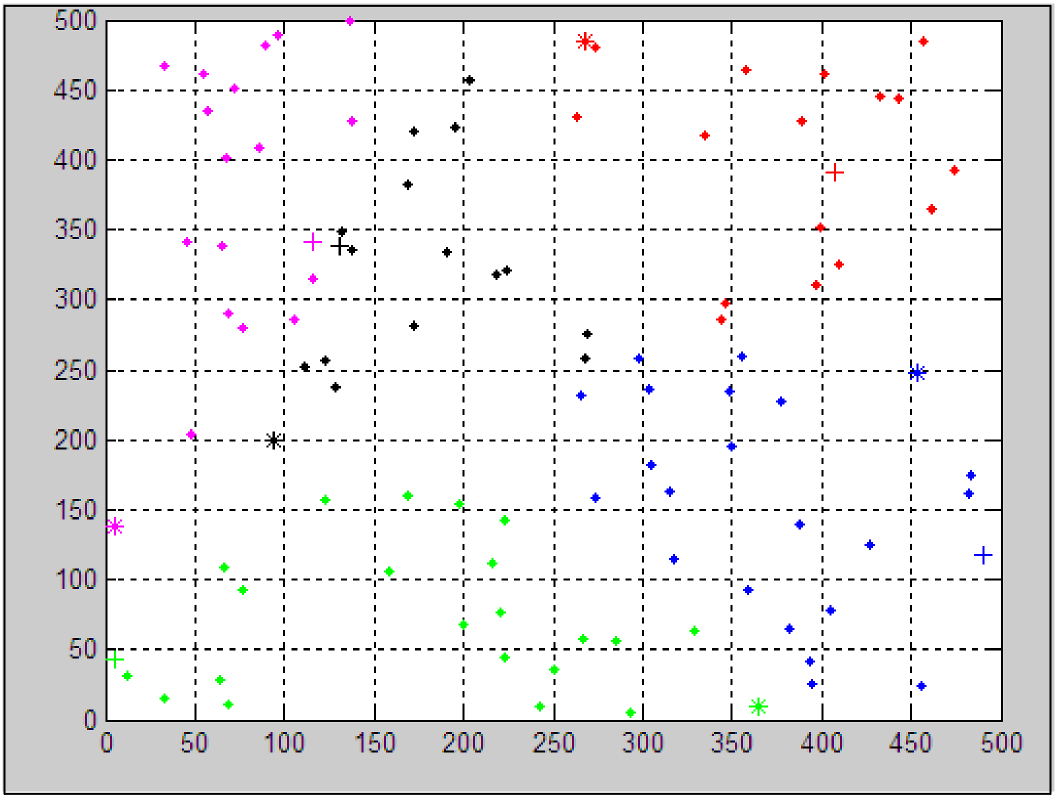

Figure 5.1: Clusters formed by Firefly-DCH

The figure 5.2 shows the clustering done by modified firefly algorithm, Firefly-DCH. Different colors represent different clusters. The * and + represents primary cluster head and secondary cluster head respectively. The clusters are formed by iteratively selecting the best solution among the possible solutions. The probability of selecting the high energy (5J) node is large as the objective function supports selection of node which has high energy. But if the high energy node is not available in the cluster then a low energy node is chosen for the selection of CH. The clusters formed by this algorithm are compact in size and the size of different clusters does not vary as the number of nodes distributed to each CH is almost uniform. This algorithm produces good network partitioning. The proposed algorithm which uses firefly algorithm prevents poor clustering due to the cooperation inherent in particles (fireflies) that can achieve global minimum of intra-cluster distance.

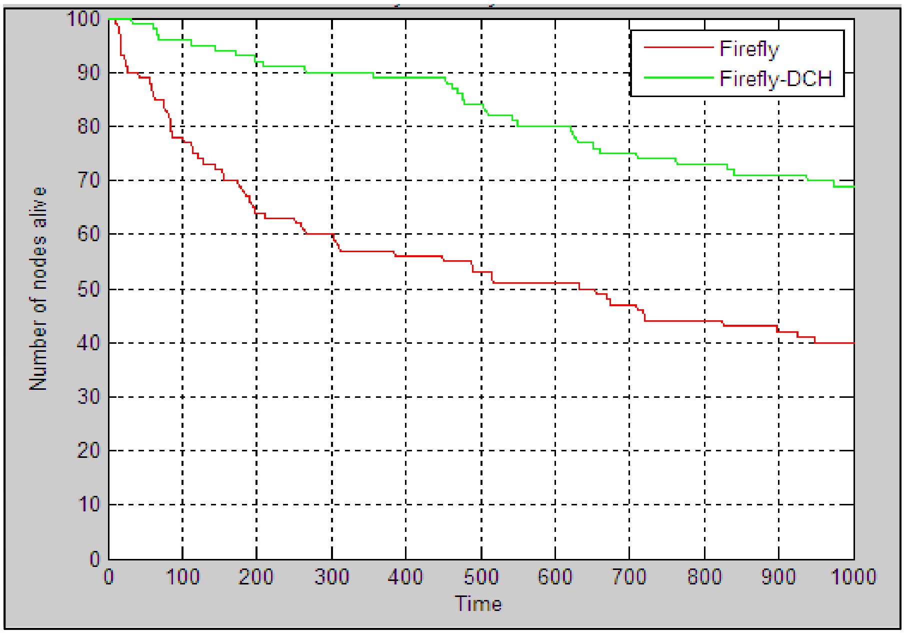

Figure 5.2: Network Lifetime (Firefly vs Firefly-DCH) with BS (250,250)

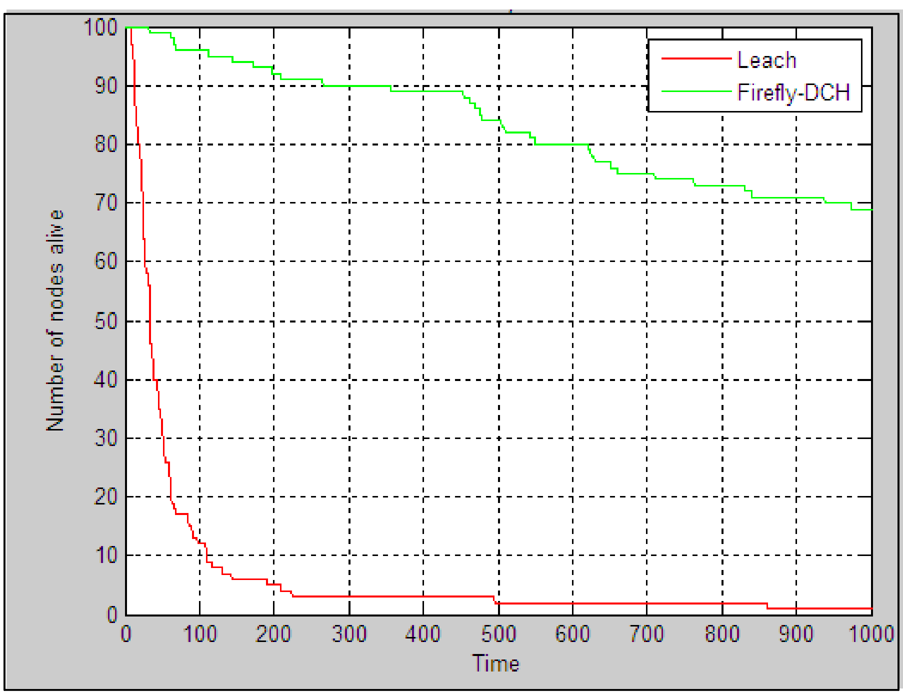

Figure 5.3: Network Lifetime (LEACH vs Firefly-DCH) with BS (250,250)

The figure 5.2 shows the network lifetime and comparison between the Firefly and Firefly-DCH. This represents the number of nodes alive over time when BS is located at (250, 250). The number of nodes alive in Firefly-DCH after 1000 rounds is more than number of nodes alive in standard firefly algorithm. The energy consumed by all the nodes in communication is reduced in Firefly-DCH due to shorter distance between the non CH node and CH. As in standard firefly algorithm there is single CH present in each cluster, so the distance between the non CH node and CH is large which leads to high dissipation of energy of the nodes.

The figure 5.3 shows the network lifetime, which clearly demonstrates that the proposed algorithm prolong the network lifetime when compared with LEACH when BS is located at (250, 250). LEACH is the homogeneous hierarchical clustering protocol in which selection of cluster head is probability based with optimal number of CHs fixed for the network. The Firefly-DCH here also performs better than the LEACH in terms of number of alive nodes. At the end of 1000 rounds there is negligible number of nodes alive in the LEACH. After 1000 rounds in Firefly-DCH, there is still optimum number of nodes alive to continue the communication between nodes and CH.

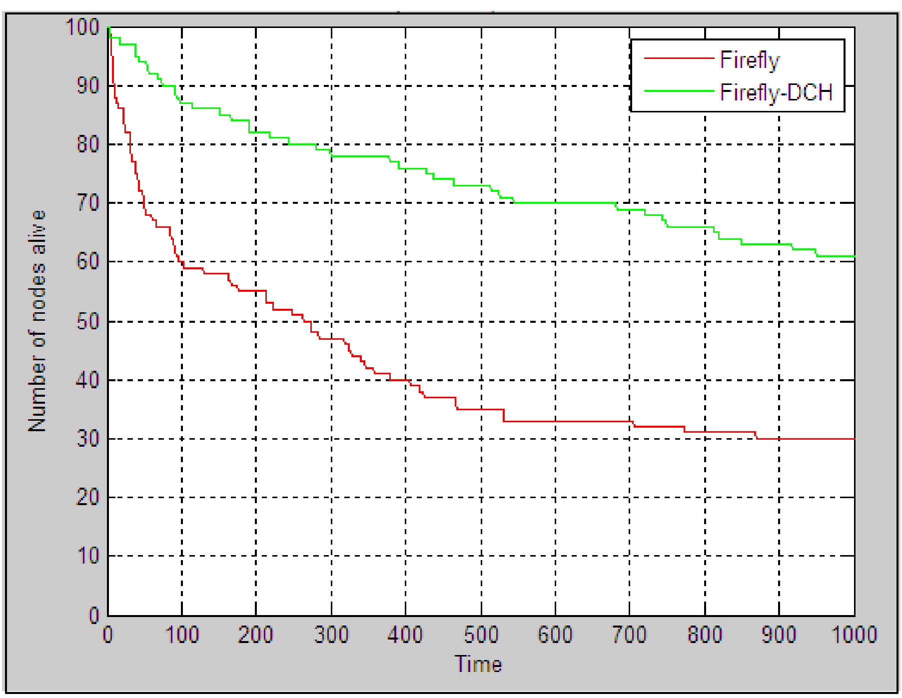

Figure 5.4: Network Lifetime (Firefly vs Firefly-DCH) with BS (250,575)

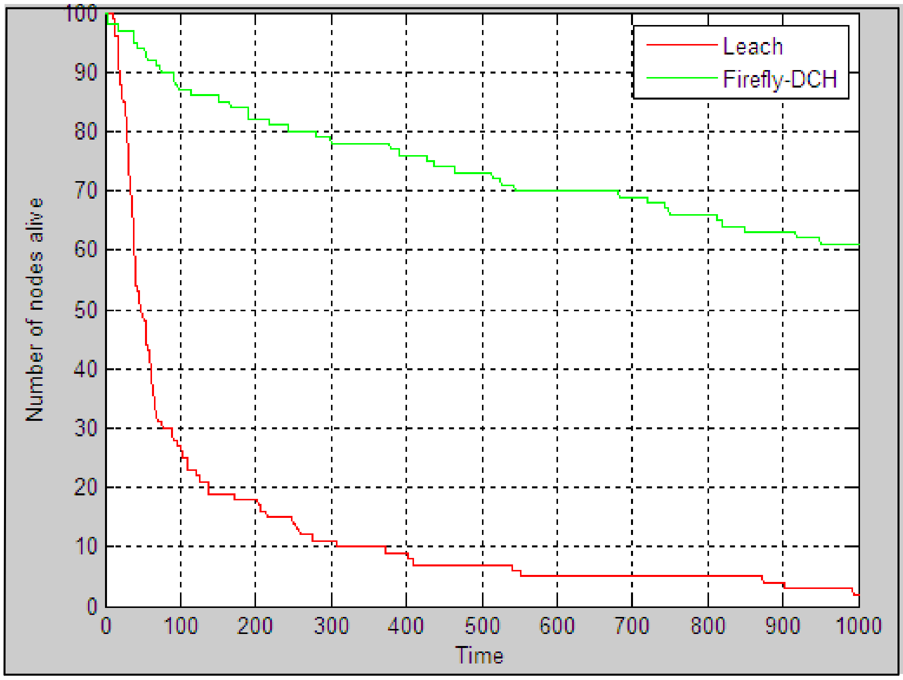

Figure 5.5: Network Lifetime (LEACH vs Firefly-DCH) with BS (250,575)

The figure 5.4 and 5.5 shows the network lifetime with BS positioned at (250,575), means when BS is 75m away from network area. In this nodes decline are more after 100 rounds because more energy is consumed by cluster head nodes in order to communicate with the BS that is located far away from the sensor field. But still our proposed algorithm performs better than standard firefly and LEACH.

Hence we have observed from the simulation results that proposed algorithm works efficiently in network for selecting CH and performing clustering. This algorithm gives better partitioning with minimum intra cluster distance and also gives cluster heads that are optimally distributed across the network. So the energy consumed by all the nodes for communication can be reduced. Thus proposed Firefly-DCH increases the network lifetime by decreasing the energy dissipation of the nodes.

CHAPTER 6

CONCLUSION AND FUTURE WORK

We dedicated our research work to WSNs and worked on finding solutions to perform efficient clustering. In this an energy aware clustering protocol is presented using Firefly-DCH algorithm. In this the objective function includes two sub functions which involves the maximum distance between the cluster head and non cluster nodes and the remaining energy of the cluster head candidates in the cluster head selection algorithm. The proposed algorithm forms better clusters by uniformly allocating the number of nodes to cluster heads and hence forms compact clusters. The Simulation results represents that the proposed Firefly-DCH algorithm gives optimum number of alive nodes which can continue the task of communication in the network. This algorithm is energy efficient and increases the lifetime of the network.

Future work in this direction would be performing clustering with the mobile nodes which changes its position after certain period of time or continuously changes its position within the network. The further work can be done to implement the multi- hop routing among the cluster head nodes to improve energy efficiency and further increase in the network lifetime. Future work can also include the implementation of hybrid optimization technique for clustering in wireless sensor network.

References

[1] J. Zhao, and A.T. Erdogan, “A novel self organizing hybrid network protocol for wireless sensor networks,” In Proceedings of the First NASA/ESA Conference on Adaptive Hardware and Systems (AHS’06), 2006.

[2] A. Milenkovi’c, C. Otto, E. Jovanov, “Wireless sensor networks for personal health monitoring: Issues and an implementation,“ Computer Communications (Special issue: Wireless Sensor Networks: Performance, Reliability, Security, and Beyond), vol.29, pp. 2521-2533, 2006.

[3] T. He, S. Krishnamurthy, J. A. Stankovic, T. Abdelzaher, L. Luo, R. Stoleru, T. Yan, and L. Gu, “Energy –Efficient Surveillance System Using Wireless sensor networks” In Mobysys’04, Boston, June 6-9, 2004.

[4] S. Fedor and M. Collier, “On the problem of energy efficiency on multi-hop vs one-hop routing in wireless sensor networks,” In Proceedings of the 21st International Conference on Advanced Information Networking and Applications Workshops (AINAW), pages 1-5, 2007.

[5] Q. Qang, H. Hassanein, and K. Xu, “A practical perspective on wireless sensor networks,” Handbook of Sensor Networks: Compact Wireless and Wired Sensing Systems, 2004.

[6] L. Subramanian and R.H. Katz, “ An architecture for building selfconfigurable systems,” In Proceedings of the 1st ACM International Symposium on Mobile Ad Hoc Networking & Computing (MobiHoc), Piscataway, NJ, USA, pages 63-73, 2000.

[7] A. Ephremides, “Energy concerns in wireless networks,” IEEE Wireless Communications Magazine, pp 48-59, August 2002.

[8] Roundy, S.P. Wright, and J. Rabaey, “A Study of Low Level Vibrations as a Power Source for Wireless Sensor Nodes”. 2002: Computer Communications.

[9] P. A. C. S. Neves, J. F. P. Fonseca, and J. J. P. C. Rodrigues, ”Simulation tools for wireless sensor networks in medicine: a comparative Study,” in Proceedings of the 1st International

Conference on Biomedical Electronics and Devices (BIODEVICES ’08), pp. 111–114, Funchal, Portugal, January 2008.

[10] D. Li, K. D. Wong, Yu H. Hu, and A. M. Sayeed, “Detection, classification and tracking of targets in distributed sensor networks,” In IEEE Signal Processing Magazine, pages 17–29, 2002.

[11] A. Mainwaring, D. Culler, J. Polastre, R. Szewczyk, and J. Anderson, “Wireless sensor networks for habitat monitoring,” In Proceedings of the 1st ACM international workshop on Wireless sensor networks and applications, pages 88–97, New York, USA, 2002.

[12] M. Liliana, C. Arboleda, N. Nidal, Comparison of clustering algorithms and protocols for wireless sensor networks, in Proceedings of IEEE CCECE/CCGEI Conference, Ottawa, Ontario, Canada, pp. 1787–1792, May 2006.

[13] H. Huang and J. Wu, A probabilistic clustering algorithm in wireless sensor networks, in Proceedings of IEEE 62nd VTC Conference, Dallas, TX, September 2005.

[14] W.R. Heinzelman, A.P. Chandrakasan, and H. Balakrishnan, Energy efficient communication protocol for wireless micro sensor networks, in Proceedings of the 33rd Hawaaian International Conference on System Sciences, January 2000.

[15] W.B. Heinzelman,A.P.Chandrakasan, andH. Balakrishnan, An application specific protocol architecture for wireless microsensor networks, IEEE Transactions on Wireless Communications, 1(4), 660–670, 2002.

[16] O. Younis and S. Fahmy, HEED: A hybrid, energy-efficient, distributed clustering approach for Ad Hoc sensor networks, IEEE Transactions on Mobile Computing, 3(4), 366–379, 2004.

[17] S. Banerjee and S. Khuller, A clustering scheme for hierarchical control in multi-hop wireless networks, in Proceedings of 20th Joint Conference of the IEEE Computer and Communications Societies (INFOCOMŠ 01), Anchorage, AK, April 2001.

[18] M. Chatterjee, S.K. Das, and D. Turgut, WCA: A weighted clustering algorithm for mobile ad hoc networks, Clustering Computing, 5, 193–204, 2002.

[19] Brucker, P., "On the complexity of clustering problems. In: M. Beckmann and H. K¨unzi (Eds.): Optimization and Operations Research". Lecture Notes in Economics and Mathematical Systems, 157, Springer-Verlag, Heidelberg, 45-54, 1978.

[20] Hussain, S., Matin, A. W., & Islam, O. (2007). Genetic Algorithm for Energy Efficient Clusters in Wireless Sensor Networks. Fourth International Conference on Information Technology (ITNG'07). doi:10.1109/ITNG.2007.97.

[21] Salehpour, A.-A., Mirmobin, B., Afzali-Kusha, A., & Mohammadi, S. (2008). An energy efficient routing protocol for cluster-based wireless sensor networks using ant colony optimization. 2008 International Conference on Innovations in Information Technology, 455–459. doi:10.1109/INNOVATIONS.2008.4781748.

[22] Kulkarni RV, Venayagamoorthy GK (2011) Particle swarm optimization in wireless-sensor networks: a brief survey. IEEE Transactions on Systems, Man, and Cybernatics-Part C. Appl Rev 41(2):262–267.

[23] Latiff NMA, Tsemenidis CC, Sheriff BS (2007) Energy-aware clustering for wireless sensor networks using particle swarm optimization, The 18th Annual IEEE International Symposium on Personal. Indoor and Mobile Radio Communications, Athens, Greece, pp 1–5, 3-7 September 2007.

[24] X. Co, H. Zhang, J. Shi, and G.Cui "Cluster heads election analysis for multi-hop wireless sensor networks based on weighted graph and particle swarm optimization". In IEEE fourth International Conference on computing, 2008, 7, 599–603.

[25] Jason Tillett, Shanchieh Jay Yang, Raghuveer Rao and Ferat" Optimal topologies for wireless sensor networks"IEEE International Conference on Personal Wireless Communications, 2004.

[26] X. S. Yang, Firefly algorithm, Nature-Inspired Metaheuristic Algorithms (2008) 79-90.

[27] W. Cheng and H. Shi, "AEEC: an Adaptive Energy Efficient Clustering Algorithm in Sensor Networks," in Industrial Electronics and Applications. ICIEA. 4th IEEE Conference on.IEEE, 2009, pp. 3950–3954.

[28] Sarma, P. N. V. S. N., & Gopi, M. (2014). Energy Efficient Clustering using Jumper Firefly Algorithm in Wireless Sensor Networks, 10(11), 525–532. doi:10.7763/IPCSIT.2014.V59.1