DATA AGGREGATION & FUSION MODELS FOR WIRELESS SENSOR NETWORK

A THESIS

SUBMITTED TO THE DELHI TECHNOLOGICAL UNIVERSITY

FOR THE AWARD OF THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

Computer Science and Engineering

SUBMITTED BY

RAJESH KUMAR YADAV

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

DELHI- 110042 (INDIA)

2018

DATA AGGREGATION & FUSION MODELS FOR WIRELESS SENSOR NETWORK

BY

RAJESH KUMAR YADAV

SUBMITTED TO THE DELHI TECHNOLOGICAL UNIVERSITY IN PARTIAL FULFILMENTS OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

COMPUTER SCIENCE AND ENGINEERING

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
DELHI- 110042 (INDIA)

2018

CERTIFICATE

This is to certify that the thesis entitled **"Data aggregation & fusion models for wireless sensor network"** being submitted by Mr. **Rajesh Kumar Yadav** (Reg. No.: 2K11/PhD/CO/09) for the award of degree of Doctor of Philosophy to the Delhi Technological University is based on the original research work carried out by him. He has worked under my supervision and has fulfilled the requirements that to our knowledge have reached the requisite standard for the submission of this thesis. It is further certified that the work embodied in this thesis has neither partially nor fully submitted to any other university or institution for the award of any degree or diploma.

Prof. Daya Gupta
(Supervisor)
Department of Computer Sc. and Engineering
Delhi Technological University, Delhi

Prof. D.K. Lobiyal
(Co-Supervisor)
School of Computer & Systems Sciences
Jawaharlal Lal Nehru University, Delhi

ACKNOWLEDGEMENT

I owe my debt and would like to express deep feelings of gratitude to accomplish the research program with the support and direction of several persons. This challenging and rewarding experience definitely helped me to grow in character as well as academically. It is a pleasant aspect that I have now the opportunity to express my gratitude for all of them.

First and foremost, I would like to thank my supervisors, Prof. Daya Gupta and Prof. D.K. Lobiyal for insightful advices, conversations and suggestions during PhD work. Their support and trust, together with genuine interest in the research subject, turned my research work into a great experience. Words cannot express my gratitude to them for their patience and support. I will be thank full to my colleagues Dr. D.K.Vishwakarma, Dr. Anil Singh Parihar, Dr. Rahul Katariya and Mr. Deva Nand for all their helps and comments throughout. I wish to pay high regards to my parents and family for their best wishes and encouragements.

Last but not least, I would like to thanks my all faculty colleague and friends who supported me during this period directly or indirectly.

Date: (Rajesh Kumar Yadav)

Place: Delhi

ABSTRACT

Wireless Sensor Networks (WSNs) is a network which formed with maximum number of sensor nodes which are positioned in an application environment to monitor the physical entities in a target area , for example temperature , pressure, environmental monitoring, water level ,industrial monitoring, health care and military application. Wireless sensor network poses two unique characteristics in comparison of traditional wireless communication system i. The limited battery power of source and ii. The redundant data, which are correlated among different sensor nodes. These two are associated with energy consumption and data traffic control. Data aggregation is a technique of gathering data from sensor nodes, eliminating redundant measurements and transmission of extracted information in an efficient way. Data fusion is also a process of combining of data from multiple sources such that the communication overhead of sending individual sensor readings to base station is reduced. The research in this thesis aims designing a model for data aggregation & fusion based on clustering, dynamic positioning and energy efficient routing for wireless sensor network.

Firstly, thesis proposed a probabilistic clustering technique for data aggregation considering differential and temporal factors for a node and cluster along with relative positions of selected cluster heads. Simulation results show that proposed technique obtains considerable improvements and allows a large stable network lifetime compared to the state of the art techniques.

Secondly, to reduce the energy consumption and redundancy,a cluster based aggregation technique using particle swarm optimization technique with discrete search space is proposed. This approach aims to minimize the intra cluster communication energy and energy loss due to cluster head and base station communication. The performance of proposed approach is compared to clustering techniques and results shows that proposed technique greatly improves over lifetime of the network.

Thirdly, a dynamic positioning based aggregation approach is proposed to reduce the redundancy and maximization of coverage area using modified artificial bee colony algorithm. In proposed approach, we have incorporated a hybrid search for balancing the exploration and exploitation. The results shows that the modifications carried out leads to enhance the lifetime and communication cost as compared to the conventional artificial bee colony algorithm.

Finally, routing based in-network data aggregation approach is proposed exploiting principles of swarm intelligence in ant colony optimization for routing and fuzzy rules to evaluate goodness of a path. Routing based data aggregation aims to maximize the number of collected data packets, while minimizing energy consumption and data gathering delay. Data collection delay should be minimum as it is key to data freshness. Due to resource constraint in wireless sensor network, minimum energy consumption can play an important role in the field of its application.

LIST OF PUBLICATIONS

[1] **Rajesh K. Yadav**, Daya Gupta, D. K. Lobiyal, "Energy Efficient Probabilistic Clustering Technique for Data Aggregation in Wireless Sensor Network," ***Wireless Personal Communications***,2017, DOI: 10.1007/s11277-017-4370-5(Pub.: Springer), [SCI Indexed].

[2] **Rajesh K. Yadav**, Daya Gupta and D.K. Lobiyal, "Energy Efficient Clustering Approach for Data Aggregation and Fusion in Wireless Sensor Networks," ***International Journal of Control Theory and Applications***, Vol.10, No. 13, pp. 241-253, 2017. [Scopus Indexed].

[3] **Rajesh K. Yadav**, Daya Gupta and D.K. Lobiyal, "Dynamic Positioning of mobile sensors using Modified Artificial Bee Colony Algorithm in a Wireless Sensor Networks," ***International Journal of Control Theory and Applications***, Vol. 10, No. 18, pp. 167-176,2017. [Scopus Indexed].

[4] Rajesh K. Yadav, Daya Gupta and D.K. Lobiyal, "Ant Colony Optimization Based Routing Approach for In Network Data Aggregation in Mobile Wireless Sensor Networks," (Communicated under review).

[5] Rajesh K. Yadav, D. Gupta and D. K. Lobiyal, "Energy efficient reactive protocol for data aggregation in Wireless Sensor Network," 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, 2016, pp. 2921-2925.

[6] R.K.Yadav et al. “A Discrete Particle Swarm Optimization Based Clustering Algorithm for Wireless Sensor Networks”Emerging ICT for Bridging the Future − Volume 2, Advances in Intelligent Systems and Computing, Vol. 338, 2015, pp 137-144.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Wireless sensor networks (WSNs) have caught the attention of the research community in recent years. The wireless sensor networks [1] are becoming a significant enabling technology in many sectors. These networks are composed of small sensor nodes that integrate sensing, wireless communication, and computation. Each sensor nodes has limited processing capability, communication bandwidth, and storage capacity. The sensor network is composed of lightweight network-enabled sensing devices, deployed in abundance to achieve network connectivity and reliability. In many cases, the deployment of the nodes is done in those areas where external power supply or an infrastructure for communication is difficult. Moreover, it is not possible to maintain nodes manually. The sensor nodes must be able to organize themselves and work unattended. The communication is a major source of energy consumption, thus sensor nodes generally self-organize into a multi-hop wireless network to avoid long-range communication [2].

1.1 Benefits and possible applications of Wireless Sensor Networks (WSNs)

The key benefits of the WSNs is that they can be implemented almost anywhere without requirement for any specific communication infrastructure. The nodes are networked in a self-organizing manner in many applications that requires unattended operations. Researchers envisage a large set of futuristic applications to be made possible with the aid of WSNs. In particular, they promise that WSNs will revolutionize the way humans interact with their physical surroundings. The following are some examples of such applications of WSN.

1.1.1 Military applications

WSN can be part of command, control, computing, communications**,** Intelligence, Surveillance, Reconnaissance and Targeting which can be utilized in military operations (e.g. battlefield surveillance, friendly / hostile forces tracking, monitoring of equipment) [2, 3, 4].

1.1.2 Environmental applications

These applications focus on collecting and reporting ambient data for general consumers. The key applications of this class are habitat monitoring ,flood detection, forest fire detection, space exploration, biological attack detection and precision agriculture [5, 6, 7, 8, 9, 10, 11, 12].The sensor nodes which can be static or mobile collect the location information and then transfer the data to central system(sink) that is the in charge of data collection.

1.1.3 Health applications

The key wireless sensor network health applications are remote monitoring of medical status, drug administration and medication intake monitoring (e.g. integrated patient monitoring, diagnostics, tracking and monitoring doctors and patients inside a hospital) [13, 14, 15, 16, 17].

1.1.4 Home applications

Wireless sensor network has the potential to provide monitoring, conservation, convenience and safety services at home. It can be utilized to establish home smart spaces by monitoring home resources and controlling equipment automatically and remotely [18, 19, 20].

1.1.5 Commercial and Industrial applications

WSN can be used for a set of advantageous services such as improvement of equipment and resources, reducing the energy cost, enhancing automatic systems, environmental monitoring in offices, inventory control, vehicle tracking and detection and traffic flow surveillance etc. [21, 22].

No doubt, sensor networks offers significant advantages in many scientific and commercial applications but limitations on sensor nodes create new challenges. Limited battery power is one of the vital constraint on sensor nodes. It is not possible to locate each node and replace its battery when it goes out of energy, especially in those areas where battery replacement is not easy. Power conservation has been well-identified important factor to prolong the network lifetime. Limited wireless communication range, susceptibility to physical damage, storage capability and low computation capacity are the other constraints of sensor nodes. As per [2] designing of sensor networks influenced by following factors:

- **Fault tolerant:** sensor nodes may be blocked, physically damaged or may go out of battery power due to these functionalities of the network should not be interrupted so there is requirement of fault tolerance.

- **Scalability:** Good scalability required for the protocol of the sensor networks because of huge number of nodes in the network.
- **Production cost:** due to the requirement of coordination among the large number of nodes, the cost of each node should be kept low to make deployment feasible.
- **Hardware constraints:** designing of sensor nodes are constrained by size, low energy consumption, low cost and capability to work unattended.
- **Topology of sensor network:** Due to node movement in mobile sensor network, failure of nodes temporary or permanent topology of network may change which needs additional nodes deployment to replace the non-working nodes.
- **Transmission media:** The nodes of network for communication purpose use wireless medium.
- **Power consumption:** Power source of sensor nodes is limited because of battery backup, so in many cases, it is not possible to replace battery of the node, which needs power management to prolong the network lifetime.

1.2 Data Aggregation & Fusion in Wireless Sensor Network

Data aggregation [23] largely determines the degree of optimization in the performance of sensor network. Data aggregation is a technique of collecting raw data from sensor nodes, eliminating redundant measurements, and extracting the information content for onward transmission. The main aim of it to gather and aggregate information in an efficient manner so that the sensor network lifetime and overall performance of the wireless sensor network is enhanced. The process ultimately diminishes the volume of accumulated information and makes in it more relevant. Data fusion is the process of combining of data derived from multiple sources such that either the resulting information is in some sense better than would be possible with individual sources or the communication over head of sending individual sensor readings to the base station is reduced. Fig. 1.1 shows how routing impacts on data aggregation. Data aggregation mechanisms in WSNs can be classified in three category: structure-based, structure-free and hybrid structure, which combines characteristics from structure-free and structure-based.

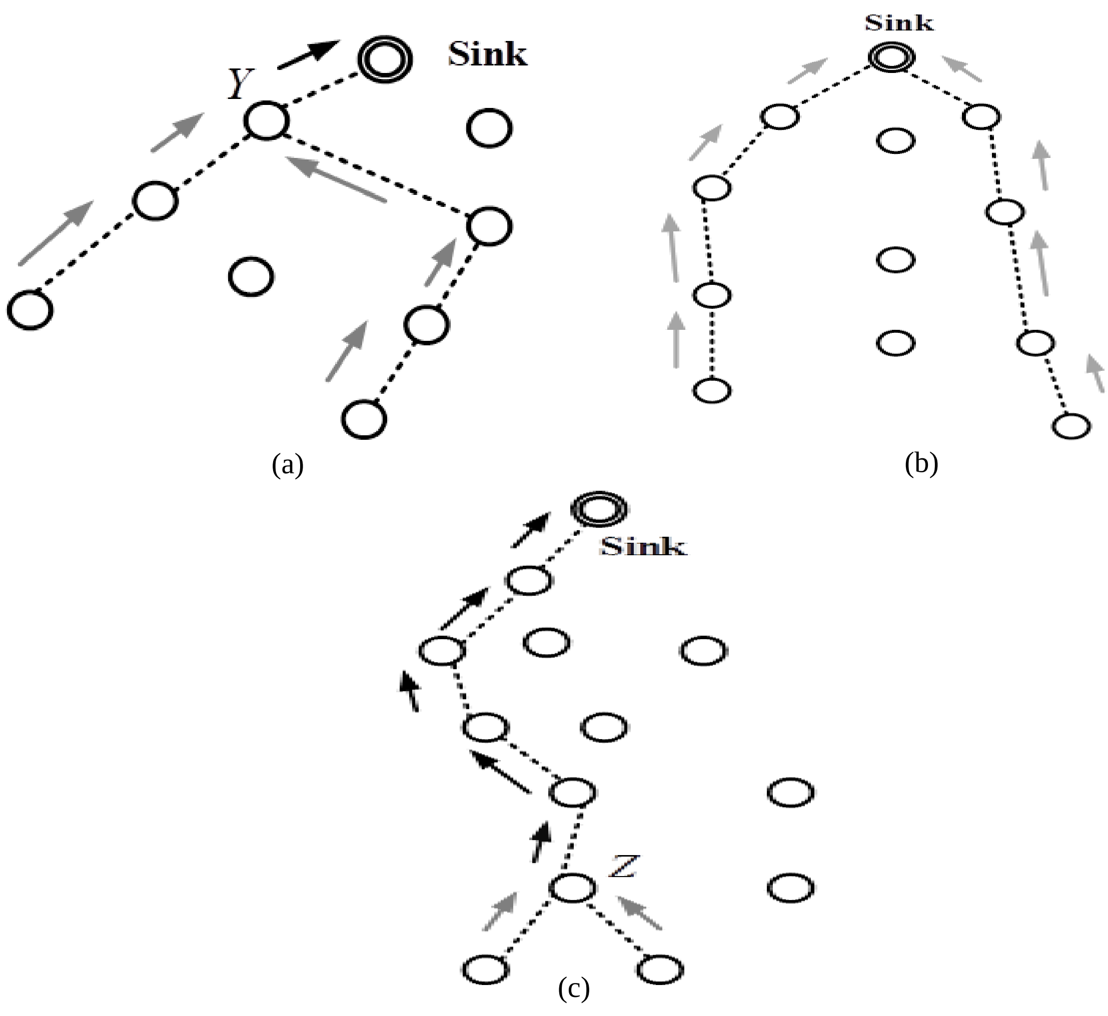

Fig. 1.1: Routing impact on data aggregation (a) Shows aggregation at node Y, (b) Not aggregated (c) Aggregated at node Z

However, WSNs are typically flat and randomly deployed. Hence, by nature, they require a structure-free mechanism. At the same time, the WSN are supposed to be deployed at a large scale, which introduces a complexity in terms of data aggregation and management. Therefore, a structure-based data aggregation defines a set of algorithms, made by the system developers, to divide the network into groups and/or levels. Then, these groups manage separately their data aggregation and offer to the system developer a reduced view of the network. Indeed, the structure-based mechanisms require an extra charge to organize the network and to maintain this organization during the network lifetime.

1.3 Key Issues in Data Aggregation

Each data aggregation technique has some issues like redundancy, delay accuracy and traffic load. Because of these issues, the performance of data aggregation gets affected. There are three main issues, which concern with data aggregation techniques for periodic information gathering in wireless sensor networks.

- Which sensor nodes should send their information?
- When should they transmit it?
- What should be the underlying routing technique to transmit sensed information?

The architecture of the sensor network also influences the performance of the data aggregation technique. The sensor network structure mainly divided into flat networks and hierarchical or cluster based networks.

1.4 Types of Data Aggregation

Based on interactions between the sensor nodes, there are four types [24] of aggregation:

1.4.1 Centralized Aggregation

In centralized aggregation scheme [25], the sensing nodes forward data to the sink by following the shortest path. Aggregation is not performed along the forwarding nodes. The fusion process is accomplished after the entire set of data arrives at the data processing center (sink). It is suitable for single-hop sensor networks. Fig. 1.2 shows the increase in the flow of traffic as the data packets reach near the sink. The darkened arrows towards the sink reflect the heavy congestion. Consequences: Increase in data traffic, Heavy congestion in the network and frequent packet drops and huge loss of information.

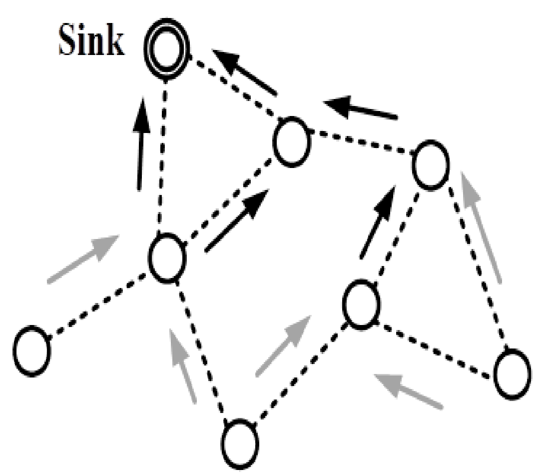

Fig. 1.2: Shows flow of data in centralized aggregation scheme

1.4.2 Tree-based Aggregation

The implementation of tree-based aggregation [23, 26] requires the formation of spanning tree over the sensor network, with sink in the root position. During the transmission of the target information, the sensors (parent nodes) wait till a period of time (aggregation time) for the data to arrive from the multiple sources. Once data reaches at the parent nodes, aggregation is performed and aggregated data forwarded to the sink through the shortest path. In Fig. 1.3, the flow of data is highlighted over the shortest path-spanning tree, thereby reducing the traffic in the sensor network. It is suitable for networks with shorter hop counts from sensors to sink, sensors having smaller sensing range and sensor networks having perfectly balanced spanning tree. Consequences: Overhead of constructing and maintaining the spanning tree, Large end-to-end delays involved in aggregation along overlapping paths, Longer aggregation time results in better fusion of data but poor response time and aggregation over shortest paths is not optimal because paths from different sources to the sink may not be merged near the target area.

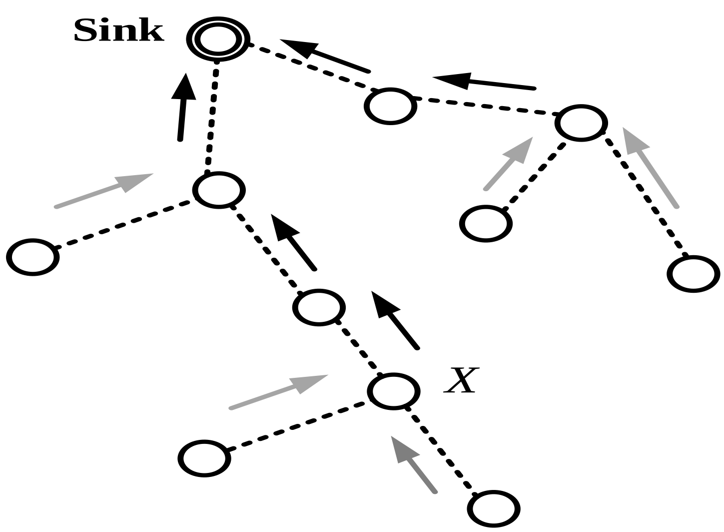

Fig. 1.3: Shows flow of data in tree-based aggregation scheme

1.4.3 Static Cluster Aggregation

In static cluster aggregation [27, 28] Fig. 1.4 the clusters are statically formed during the initial network setup phase. Once formed, cluster structure remains unchanged (static). Each cluster thereafter appoints a cluster head and rest of the members sends data to it. The cluster head performs aggregation and reports to the sink. It is suitable for Networks with less or no mobility and Area monitoring applications (recording earthquake, temperature, humidity, etc.). Consequences: Static cluster often fails to encapsulate the requirements of the dynamic applications and not suitable for capturing the mobility of the sensor networks.

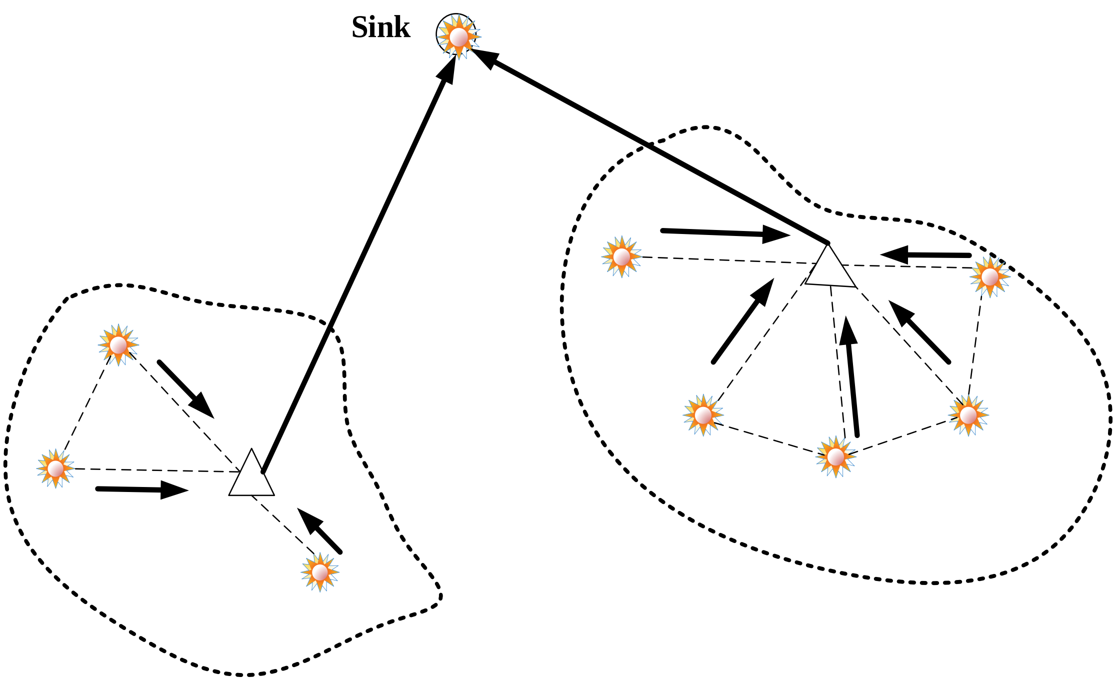

Fig. 1.4: Shows flow of data in static cluster aggregation scheme

1.4.4 Dynamic Cluster Aggregation

In dynamic-cluster aggregation scheme [29, 30, 31, 32, 33, 34], clusters are formed dynamically by sensing environmental parameters or target information. Moreover, the initial cluster structure is updated periodically (adaptive clusters). Multiple sensor nodes collaborate with each other to elect a cluster head. The elected cluster head receives data from its members and transmits an aggregated data packet to the sink. It is suitable for sensors with variable sensing ranges, Large sized sensor networks with increased hop count between source and sink, Mobile wireless sensor networks and Wide range of dynamic applications: forest fire supervision, wildlife monitoring, target tracking, etc. Consequences: Better energy utilization and load balancing in the network and exhibits higher degree of collaboration among the sensor nodes. The correlation in wireless sensor network and exploiting the correlation of sensory data have been studied in [35].In [36] hybrid data collection approach have been proposed that dynamically switches between event driven data gathering and time driven data gathering. They have suggested that the sensor nodes that seem to detect an event of interest in near future, as well as those nodes detecting the events, became engaged in the time driven data reporting mechanism.

Data aggregation collects data from different sources and then combines them using an aggregation function. It has the potential to decrease the number of transmissions by combining the set of data into a single packet to transmit. Decreasing number of transmissions reduces the communication costs, bandwidth utilization, network congestion, energy consumption and network delay in wireless sensor network. According to these advantages and due to node resource constraints and existing drawbacks this technique is highly desirable to be used in WSN. Data aggregation and fusion has three elements: Data representation, aggregation function and routing. Data representation focuses on semantic and syntactic attributes. Data aggregation function is a method that is utilized to combines data samples which depends on end user query and applications. Average, maximum, minimum, count median, rank, standard deviation, variance, sum and regression are key aggregation functions. Data aggregation routing is a mechanism of establishing communication paths between the source nodes and the sink to forward the data. It usually includes path discovery, setting up either single or multi hop data forwarding and path maintenance.

The performance of data aggregation protocols can be measured by different metrics, such as energy cost, data forwarding delay, and end-to-end loss rate. Real time monitoring applications aim to retain a “current” view of the monitored area and require effective aggregation with minimal forwarding delay. For example, in battlefield surveillance, sensor nodes are deployed to detect and track moving military targets such as tanks and other vehicles [37] in which the data collected from the sensor network should remain as current as possible. In data gathering applications huge amount of communications is reduced by in-network aggregation achieving maximum lifetime of the network [38] .In general low data forwarding delay is important in any application requiring response to external phenomenon in real time [39]. In some cases 100 percent, reliable delivery is expected, while in some cases moderate levels of loss can be tolerated.

In an effective aggregation, timing also plays an important role, sometimes aggregation requires that sensed data be delayed at intermediate nodes, while waiting other data to be received. Data collection in sensor networks can be done by using either unicast or broadcast [40, 41, 42] a unicast packet is received and processed by the intended recipient only. A huge amount of work has been done and evaluated over routing schemes in sensor network [43, 44, 45, 46, 47, 48]. Different technologies for support, example time synchronization and other quality of service have been extensively studied [49, 50, 51]. For the conservation of energy, different technologies have also been proposed in sensor networks. Data aggregation combines data from different nodes to suppress redundancy and traffic volume for reduction of energy consumption [52, 53, 54]. Broadcast based data collection takes advantage of the inherent redundancy of the broadcast medium in sensor networks for reliability [40, 55]. When packets are transmitted, using broadcast based aggregation mechanism then all nodes receives the packets those are in the range of transmission. However, the drawback of broadcast mechanism is that same packet is received and aggregated by multiple nodes, which may produce wrong results for duplicate sensitive aggregation function.

The evolutionary algorithms are optimization algorithm of the meta-heuristic or stochastic nature. Thus, a global optimum is not guaranteed. However, these algorithms give near-optimal solutions. The evolutionary algorithms are used in complex problems/ NP-Hard problem, where traditional algorithms fail. There is various evolutionary algorithms available in the literature like Particle swarm optimization (PSO), Artificial Bee Colony algorithm (ABC), Ant Colony Optimization algorithm (ACO). For implementing individual sensor nodes in WSN the better optimization approaches which require reasonable resources to produce better results. Particle swarm optimization (PSO) is an optimization technique in which natural species social behaviors are considered for the purpose of computation. It is a swarm intelligence technique based on population that performs optimization process with the objective of optimizing a fitness function. It has the advantage of solutions with better quality and higher efficiency in computation [56, 57, 58]. The ABC algorithm mimics the foraging behavior of honeybees. The food sources in the ABC algorithm denote a solution in the search space. The fitness of a food source is determined by its nectar value. In ABC algorithm a colony of bees is used as the swarm, the swarm of bees is made up of three types of bees namely the employee bee, the onlooker bee and the scout bee [59, 60]. Ant colony [61, 62, 63] is a standard solution for finding optimal path from source to destination that is proposed by authors for solving optimization problems such as traveling salesman problem (TSP) with multi-agents. The fuzzy theory [64, 65] is also very useful to determine the optimal paths between source and base station for wireless sensor network (WSN).

1.5 Major Challenges

Based on the analysis of the existing state-of-art algorithms for cluster based data aggregation and fusion in wireless sensor networks.

- In literature most of the work done based on probabilistic clustering techniques but these techniques do not take into account the relative positions of selected cluster heads; and uses fixed inter cluster head election latency and also not considered differential & Temporal factors for a node and cluster, which gives rise unbalanced clustering and wastage of energy in inter cluster communication.
- Partitioning the network into optimal number of clusters and selecting an optimal set of nodes as CHs is N-P hard problem.
- Determining the optimal positions of the sensors to maximize area coverage and it becomes harder when sensor nodes are mobile.

- Power consumption remains one of the critical issue for routing in mobile adhoc sensor networks.

1.6 Problem Statement

A wireless sensor network is designed to gather the information through the area and the sensed information must be transmitted to a central node i.e. a base station or sink. The technique through which data is gathered and transmitted to sink node through a network has a great impact on lifetime of the network and energy consumption. Sensor nodes deployed in the area of interest may send sensed data directly or indirectly to sink node. In both direct mode sensor has to upload information to base station or sink using one hop wireless communication, while in indirect mode information transmitted by sensors using multi-hop wireless communication but due to short communication range of sensor nodes base station node communicate with limited number of sensor nodes. In WSN, each sensor node has limited storage capacity so some nodes may fail to receive or transmit information further to base station or sink node. Positioning of sensor nodes also affects the overall performance of the sensor networks so by considering these the statement of problem addressed in this thesis that how to deploy and reduce the transmissions rate of the sensors in order to optimize network lifetime.

1.7 Research Goals

The main goal of the thesis is to study the performance of different data aggregation and fusion protocols for periodic data collection in sensor networks. The cluster-based aggregation has three fundamental issues are addressed in the thesis. The first issue deals with the balanced clustering and cluster head selection. The second issue deals with the determining of the optimal positions of the sensors if sensor nodes are mobile. The third issue deals with how packets are transmitted to sink node.

- First part of the thesis explored issues of probabilistic clustering and cluster head selection in which timing control has a great impact on aggregation efficiency and data forwarding delay.
- The second part of the thesis examines the performance of coverage maximization and redundancy minimization with data aggregation based on node deployment.
- The third part of the thesis examines routing based aggregation that minimizes the number of packet transmissions, increases network lifetime by reducing routing overhead and network energy consumption. It also reduces routing delay by establishing minimum hop count paths.

1.8 Contributions of the Thesis

This thesis investigates protocol design and performance issues related to periodic data aggregation & fusion for wireless sensor network. In particular, probabilistic cluster based aggregation, dynamic deployment based aggregation and routing based aggregation are investigated.

First, we have developed probabilistic clustering technique considering relative positions of selected cluster to ensure more or less optimal number of clusters and well distributed CHs across the network in each round with two generic approaches to integrate the proposed technique to existing algorithms.We also used differential temporal and energy based aggregation to optimize lifetime of the wireless sensor network by using optimal differential inter cluster head election latency determination and optimal differential energy based cluster head selection approach.Simulation results show that proposed technique obtains considerable improvement and allows a large stable network.

"This work has been published in Wireless Personal Communications*, DOI: 10.1007/s11277-017-4370-5, 2017(Pub.:* ***Springer)****, [SCI Indexed].***"**

Second, we proposed a PSO based solution to clustering problem in which a new operator is defined to make it work with discrete search space. Here aim is to minimize the intra-cluster communication energy and energy loss due to cluster head and base station communication. Simulation results show a considerable increment in Network lifetime.

"This work has been published in International Journal of Control Theory and Applications*, Vol.10, No. 13, pp. 241-253,2017 [Scopus Indexed]***"**

Third, We have developed an algorithm for dynamic positioning of sensor nodes in the area of interest using artificial bee colony algorithm with the search for new solutions be guided along the global best solution found till that instant instead of random searching around an old solution. Another shortcoming of the ABC is also addressed which balances exploration and exploitation phase. In proposed work, we have incorporated a memetic search for balancing the exploration and exploitation along with another possible scope of improvement in the functioning of ABC algorithm is to use the fittest food sources to generate new food sources using a cross over operator.

"This work has been published in International Journal of Control Theory and Applications*, Vol. 10, No. 18, pp. 167-176,2017 [Scopus Indexed].***"**

Lastly, we have proposed a routing approach, which exploit the principles of swarm intelligence inherent in ant colony optimization (ACO), and the goodness of a path is evaluated by a set of fuzzy rules. Simulation results show that the performance of the proposed protocol is improved in comparison of a standard state of the art routing protocol.

"This work is Communicated to ***Ad Hoc & Sensor Wireless Networks (under review)****"*

1.9 Thesis Organization

The remainder of the thesis is organized as follows: **Chapter 2**: review prior work on data aggregation and fusion. It also gives the highlights of research gap in the concerned area and based on the research gap, the objectives formulated and explained. **Chapter 3**: proposes an algorithm based on the probabilistic clustering approach, which maintains the requirement of clusters of almost same size and the optimal percentage of CHs in each round. It also includes differential temporal and energy based aggregation to optimize lifetime of the wireless sensor network with experimental results and analysis. **Chapter 4:** considers data aggregation based on particle swarm optimization with discrete search space and hybrid clustering protocol (HCP) for multilevel heterogeneous wireless sensor networks with experimental evaluation of proposed algorithm. **Chapter 5**: presents new coverage maximization approach using dynamic deployment of mobile sensors based on modified -ABC and compares the performance of the proposed protocol with conventional ABC. **Chapter 6:** describes Ant Colony Optimization Based Routing Approach for In-Network Data Aggregation. **Chapter 7**: summarizes the thesis and outlines future directions.

CHAPTER 2

LITERATURE REVIEW

In wireless sensor networks, Batteries of the sensor nodes generally not replaced so energy is the most careful analysis parameter. Power consumption is an important issue in sensor networks. Communication is considerably more energy consuming than computation as it has been shown in past research. Transmitting a single bit information may cost hundreds or even thousands of times as much energy as executing a single instruction [66, 67]. Based on a basic observation data aggregation has been proposed to improve energy efficiency in sensor networks [52, 68, 69].The main idea is to perform in-network processing to reduce communication cost. If sensor nodes are densely deployed then sensed data produced by neighboring nodes may be highly correlated and redundant. Particularly in event detection applications, some event may be detected and reported by many nodes. Transmitting of data consumes more energy in comparison of processing; energy consumption can be improved by reducing the amount of transmissions through in-network processing or data aggregation [32]. In [70] authors have identified three basic parameters in data aggregation: efficient networking protocols, effective aggregation functions and efficient ways of representing data. A well-designed routing algorithm is one of the best ingredients for data aggregation. Traditional routing algorithms mainly focused on the shortest path routing mechanism. In sensor networks, however, data-centric routing is adopted mostly because it is based on the content of data packets, which also promotes data aggregation.

In data collection applications such as habitat, monitoring and environment monitoring nodes collect data after some interval and collected data transmitted to sink node [6, 9, 12]. Aggregation is performed as packets are being forwarded to sink. In periodic collection aggregation, mechanism generally based on network structure such as trees and clusters. Nodes those are participating in the task are relatively stable, which may increase cost of construct and maintain the structure. While in event-based application, source nodes change dynamically according to requirement or situation [71, 72]. The overhead of building and maintaining tree structure may be costly. A structure free protocol also has been proposed to avoid the maintenance cost of the underlying structure [73]. Multipath routing based aggregation protocols are also have been proposed [40, 41, 55, 74, 75]. Unicast based data aggregation protocols, which are further, classified as cluster based, tree-based, structure free, semi-structured and data aggregation based on other structures. In structure-based protocols, nodes generally form a structure such as tree or cluster in which aggregation is performed at intermediate nodes as data is transmitted to the sink over the structure. There are some structure free data aggregation protocols have also been proposed for event-based applications with dynamic source nodes [73, 69]. Aggregation protocols for periodic data collection are usually based on network structure such as trees and clusters.

2.1 Tree-Based Aggregation

Several works [46, 41, 76, 77, 78, 79] have used the tree-based structure as a solution for efficient data aggregation in WSN. The works presented in [76] and [77] are based on the improvement of the direct diffusion (DD) protocol [46] which is one of the well-known protocols in WSN. It is a data-centric protocol, where all the communication is for named data, which are named using attribute-value pairs. After the data naming, the sink broadcasts a message (called interest) describing its desired data to its neighboring nodes. This broadcast that passes throughout the network sets up gradient to indicate the back route of the collected data. By sending the broadcast, the sensor node matches the properties of the measured data with the interest of the sink and then gives its data, if there is any correspondence. According to [77], in the data aggregation process applied with DD, data obtained by the source nodes aggregated locally after a fixed length of time, this only reduces the data flows from the source, rather than prohibiting the direct forwarding in the central nodes. Thus, the true advantages of data aggregation applied in monitoring type WSNs are not fully realized. In authors proposed an improved data aggregation mechanism based on cascading timeouts principles, with the aim of reducing redundancy caused by large amounts of forwarding. The accuracy and Real-time properties of the data are guaranteed by the introduction of cascading timeouts, which also balances data flows on different transmission paths. In [76] even though source nodes are near to the sink node, many other unnecessary nodes in the networks are involved to propagate interests and setup gradients to the whole network. Due to this, DD generates unnecessary traffic during data transmission. It also achieves energy inefficient data aggregation because sources do not know where to forward data for aggregation. In DD, data are aggregated only by chance if the gradients are established as a common path for all sources nodes. As a result, many nodes involved in aggregating data are energy inefficient. A well-known aggregation mechanism TAG [41]: a Tiny Aggregation service for ad-hoc sensor networks. It is a data-centric protocol, which is based on a tree structure aggregation and is specifically designed for monitoring applications. This means that all nodes should produce relevant information periodically. The implementation of TAG consists of two main phases: (1) the distribution phase, where queries are disseminated to the sensors and (2) the collection phase, where the aggregated sensor readings are routed up the aggregation tree. The distribution phase passes by a broadcast message from the sink in order to organize the sensor nodes into a tree. As for most tree-based schemes, TAG may be inefficient in case of dynamic topologies or link/device failures. The trees are particularly sensitive to failures at intermediate nodes as the related sub tree may become disconnected. In [42] authors have studied the impact of timing in periodic tree based aggregation. In [71, 72] authors have proposed dynamic convoy tree based collaboration (DCTC),here convoy tree is tree structure which includes only those nodes surrounding the target and it changes dynamically as the target moves around. The information is collected by the root node of the convoy tree from the source nodes then it generates a report and forwards it to sink node. However, the main challenge of this approach is how to reconfigure tree in an efficient manner as target moves. Some new nodes may join the existing tree and some existing nodes may need to be quit. Due to movement of the target root may have to be replaced also.

2.2 Cluster Based Aggregation

In cluster-based data aggregation, nodes are grouped into clusters, with one cluster head (CH) for each cluster. Cluster members(CMs) may send data to their cluster heads via single hop or multi-hop communication, cluster head aggregates received data and forward the result to sink node (base station) via single hop or multi-hop communication. In Fig. 2.1 cluster members communicate with their cluster heads (CHs) directly and the cluster heads communicate with sink node (base station) directly. Important design issues in clustering systems include energy consumption balancing, cluster head placement, and determining the optimal number of cluster heads.

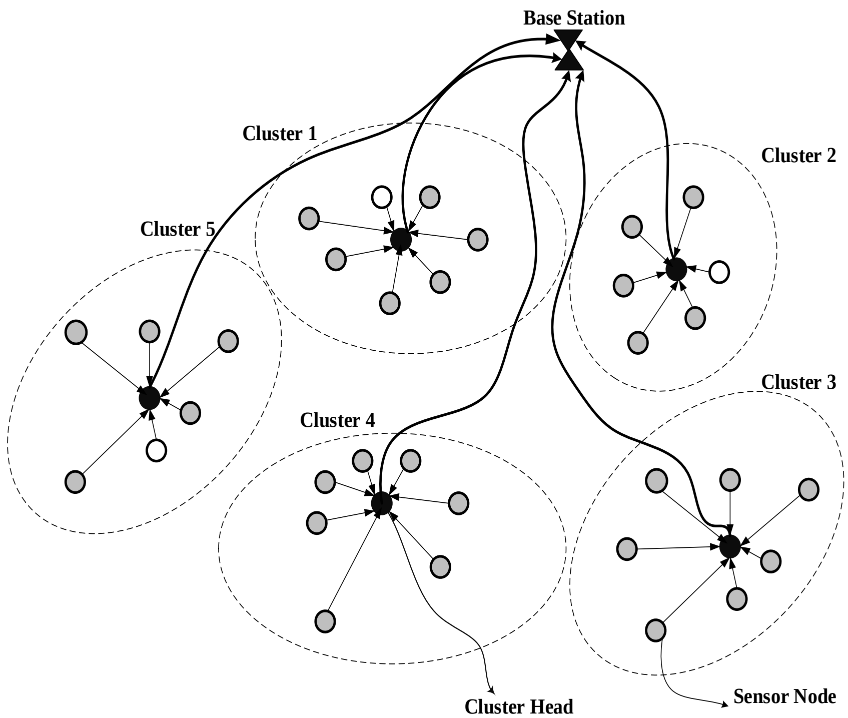

Fig. 2.1: Cluster-based Data Aggregation

Low energy adaptive clustering hierarchy (LEACH) [34] is a widely known example of this type of algorithms. It protects from battery reduction and stability in nodes energy consumption. However, it wastes energy during Cluster Head (CH) selection phase & uses a huge amount of energy when CH is at large distance from the sink. Moreover, it doesn’t guarantee of good CH distribution. It is a cluster based data aggregation protocol in which nodes are assumed to be able to adjust their power level to achieve different transmission ranges. Based on randomization function a predetermined fraction of the nodes elects themselves as cluster heads and broad cast a message to the whole network. Other nodes decide which cluster they should join based on the signal level of the messages that they received. Each cluster members are assigned different time slot based on TDMA mechanism by the respective cluster heads. CHs receives data from each member of the cluster then aggregates the received data before sending to sink node via CDMA to avoid interference between clusters. LEACH avoids excessive energy consumption of the cluster heads by rotating cluster head positions among nodes. It has two phases: setup and study state phase. The former focuses on cluster formation, whereas latter transmits data packets from source nodes to sink. The cluster heads (CHs) are periodically selected based on a distributed random algorithm in which each cluster member (CM) may become a CH for a particular round according to probability value (p). The probability value allows a cluster member to become a CH for 1/p round.The source node collects and transmit data packets to CHs using Time division multiple access (TDMA) [80] to avoid intra cluster collisions. Code division multiple access (CDMA) is used by CHs to avoid inter-cluster interference. LEACH uses periodical CH selection to replace low battery CHs with new ones would increase energy consumption. It establishes one hop inter and or intra cluster links between the sink, CHs and cluster members to forward data which is not suitable for large network. It has non-balance cluster heads distribution and uncertainty in cluster count and size. It also uses distributed way for cluster head election and cluster formation, which may lead poor cluster head placement. To resolve these draw backs many protocols have been proposed. TLLEACH [81] is a two level LEACH resolves inter/intra cluster single hop communications by establishing a two-level clustered infrastructure, they are called primary and secondary clusters. MRLEACH (Multi-hop routing with LEACH) [82] provides multi-hop paths between the CHs to transmit data packets to the sink. M-LEACH [83] it is suitable for large size network but it suffers from hotspot and limited scalability.

As the utility of WSNs is being realized, in hindsight the various challenges related to sensor networks are coming to the fore, one such issue is the dynamic deployment of the sensor nodes in a monitoring area. ELEACH (energy aware LEACH) [84] this protocol initially selects CHs similar to LEACH randomly and then utilizes residual energy at each node to selects the CHs for next round. In Low energy adaptive clustering hierarchy-centralized (LEACH-C) [37] the base station initiates centralized algorithm to elect the CHs according to their location information. It forms better-balanced Clusters. However, it wastes energy to attain global information & not robust. The sensor nodes forward the required clustering information such as location, residual energy and connectivity degree to the sink during set up phase. The sink node proactively forms a set of balance clusters in terms of energy, coverage and connectivity then allocates the role of CHs or CMs to the nodes. However, collection of clustering information by the sink to form clusters is a drawback of LEACH-C. V-LEACH (Vice Cluster head LEACH) [85] is a protocol in which vice cluster head is also selected in each cluster to handle cluster communication and computation in case of CH failure. CHs may go out of energy very fast because of heavy load in comparison of cluster members in this situation vice cluster head would stay in the cluster to takes the responsibility of CH if it fails. LEACH-FL (LEACH fuzzy logic) [86] in this approach fuzzy logic based on three parameters: residual energy, density, and distance from sink node utilizes to select cluster heads (CHs).it has potential to reduce energy consumption in the selection of CH and consequently enhances the network lifetime. In authors have proposed two hybrid clustering based data aggregation in target tracking application of wireless sensor networks. Multiple clustering technique is applied simultaneously in a single network depending on network environment. Low energy consumption and high packet transmission ratio are achieved because of clustering based on the velocity and number of the targets [87].

Hybrid energy-efficient distributed clustering approach (HEED) [88] is another typical clustering algorithm in WSNs and the cluster head election strategy of it is directly related to node's residual energy. However, HEED did not take into account of network structure that causes the imbalance of energy consumption about the cluster. It considers two parameters to select cluster heads and form the clusters: residual energy and node connectivity degree. In this algorithm, clusters are formed by selecting the nodes that have sufficient energy and are able to dominate a specific number of nodes as CHs.it would result in forming well-distributed clusters in the network. It performs better than LEACH in terms of energy consumption due to minimization of the number of re-clustering rotations. The network consumer manages reforming of clusters according to a finite iteration number. In addition to this lifetime of CHs are increased because of the selection based on energy-aware approach. On the other hand, in LEACH cluster heads have a higher chance to fail as they are selected based on probability value and without considering energy level. However finding an efficient number of clustering rotation is challenging issue in HEED. If the clusters are formed frequently then resource consumption increase and if clusters are formed less frequently then bottleneck and cluster head failure increases.

Power-efficient gathering in sensor information system (PEGASIS) [89] protocol is an enhancement over LEACH. In order to extend network lifetime, nodes are organized to form a chain, so that they need to communicate only with their closest neighbors. This reduces the power required to transmit data per round as the power draining is spread uniformly over all nodes. A greedy chain construction scheme is proposed in this algorithm, which needs global knowledge of the network at each node. Based on residual energy and location information, it selects a set of nodes as leader nodes to collect, aggregate and transmits data packets. Greedy algorithm is to forward the data packets from source nodes to next hop nodes if they are closer to leader nodes. The leader nodes are responsible for transmitting results to sink. The nodes construct a new chain if leader node fails. The difference of PEGASIS and LEACH is that the sensor nodes do not need to pay frequently clustering cost to re-cluster the network. It out performs LEACH in terms of network life time but over head of leader selection is increased in PEGASIS when the network works over long period. Data collection delay is also increased in PEGASIS due to multi hop transmissions from source node to sink. Hierarchical PEGASIS is an improvement over PEGASIS by the same author aiming to solve the delay drawback. It reduces delay using parallel transmissions from source nodes to the sink. H-PEGASIS uses two techniques to provide parallel communications: 1. Signal coding, in which the nodes construct a tree of chains, which is rooted in the sink. Data packets are transmitted by the node to their leaders in parallel where each level of transmission is coded by CDMA that allows collision free parallel communications. 2. Transmitting spatially separated data, which allows the nodes that physically close to each other to transmit data packets to the leaders at each round. Unequal clustering size (UCS) is proposed for providing uniform energy dissipation it solves the problem by varying the size of the clusters with respect to the relay load it is extended to homogeneous networks in which packets are forwarded to the sink node over a backbone formed by the cluster heads. In [90] authors have considered heterogeneous network in which sensor nodes may use single hop or multiple hop communication for the communication with their cluster heads while cluster heads communicate directly to sink node with long range transmissions.in single hop communication nodes further away from sink node spend more energy on transmission, however in multi-hop communication the nodes closest to the sink have higher energy communication because of their relaying burden. In [91] authors explored how to determine the optimal number of cluster heads to minimize total energy consumption in a homogeneous network. There are some applications demand real time delivery, which needs tradeoff between waiting delay and aggregation accuracy. Cluster heads generally wait for a certain period of time to aggregate the packets it has received and sends result to sink node but in this case packets received after transmitting of aggregated result may not be aggregated in that particular round. DSFC (Delay sensitive feedback control) [92] proposed a mechanism to solve the trade-off in cluster-based aggregation. In DSFC, the ratio of the number of packets aggregated by the cluster head to the total number of packets received by the cluster head for a round is defined as the aggregation scale. Means difference between the aggregate it transmitted and the aggregate of all received packets for that round. The error is then used to adjust the required aggregation scale for next round. If the error is bigger than a specific value then aggregation scale is increased [93, 94, 95, 96, 74, 97, 98, 99].

TEEN (Threshold sensitive energy efficient sensor network protocol) [100] utilizes a reactive-DC (data centric) approach to cluster the network in which data packets are reported in a time sensitive manner. It restrains needless data transmission by buffer mechanism to reduce data stream in the networks, sequentially can save energy of sensor nodes so as to prolong the lifetime of networks. The source nodes are grouped into a set of clusters based on their available data samples. They are assigned by two parameters: Hard threshold (HT) and soft threshold (ST). The hard threshold is a threshold value of sensed data that shows new type or value data. The soft threshold is acceptable change rate of similar data samples. It leads to report data in the case of sudden changes. Data samples are transmitted to a cluster head based on threshold value, if it is beyond HT and /or greater than soft threshold.Threshold values are also considered by the CHs to report collected and aggregated data to sink. It does not allow all source nodes to transmit data according to changes that are interesting for the data consumer. Data collection delays are reduced because of few source nodes try to access the wireless channels to forward data packets. However, it is not suitable for periodic data collection as it returns no result until data values are changed and threshold is reached. Adaptive periodic threshold sensitive energy efficient sensor network protocol (APTEEN) [101] resolves the existing drawback of TEEN. It is able to support three types of queries: a. historically based on past data analysis, b. one time and c. persistent by continuous monitoring the environmental events over a period. In this approach, each cluster head (CH) sends a message containing four parameters attributes, thresholds, scheduling metric and count time. Attributes show the elements of data in which data consumer interested. Scheduling metric assigns a time slot to each CM to provide intra cluster scheduling, while count time determines maximum allowed period between two data reports, source nodes re-sense or retransmit data if CT counter is expired. Hence, CT enhances data collection flexibility when data consumers need periodic data gathering. Assigning the threshold values to the nodes is a challenging issue for both TEEN and APTEEN algorithms. Threshold value allocation also consumes network resources if the values are frequently transmitted from sink to the CHs especially in a dense network. In [102] authors proposed a novel cluster-based data aggregation protocol which divides network into grids with unequal size, the grid further away from BS has bigger size and more nodes. The CHs rotation is performed in each grid. Although the CHs in the grids that are further away from BS consume more energy in each round, these grids have more nodes to participate in CHs rotation and share energy load to balance energy dissipation. Some of the researches performed for improving aggregation in sensor network used specialized branch of computational intelligence, for instance entropy and information theory, wavelet theory, Brownian motion etc. [103]. In significance research in [28] the authors have provided an energy efficient and loss less data aggregation scheme for sensor network scenario. In [30] Brownian Motion is applied to model the data gathering process, which reduces redundancy in the sensed data by performing filtration at sensor node of network hierarchy. Subsequently entropy based aggregation is performed at sensors, followed by wavelet-based data aggregation at cluster head resulting in minimized number of packet transmissions. In [98] authors have proposed data density correlation degree based clustering method for data aggregation Power efficient data gathering and aggregation protocol (PEDAP) [104], a type of minimum spanning tree is constructed to minimize energy consumption in WSNs, but the residual energy of the sensor node is neglected. Energy efficient clustering algorithm (EECA) for data aggregation has been proposed [99] by considering the node 's residual energy as well as the average distance to its neighbors, the cluster head is selected and an aggregation tree is constructed to save energy on communication. SEP (Stable Election Protocol) [105] it is a heterogeneous proactive protocol, which consumes energy from the nodes having high energy and increase the stability period and lifetime of the network. In [106] authors proposed to utilize mobility for joint energy replenishment and data gathering. A multifunctional mobile entity, called SenCar was employed, not only as a mobile data collector roaming over the field to gather data via short-range communication but also as an energy transporter that charges static sensors on its migration tour via wireless energy transmissions. In [107] Author proposed energy efficient adaptive multipath routing technique to reduce overhead and efficiently utilizes the energy availability. Grid based aggregation protocols also found in the paper [108] in which entire target area is divided into number of grids and inside each grid one node is selected as data aggregator [109].Authors in the paper [110] suggested that hierarchical clustering save energy in tree structured wireless sensor network. It conserves communication bandwidth and stabilizes network topologies. The authors in [111] have shown that particle swarm optimization out performs over probabilistic clustering algorithms and increase the network span and overall throughput. The Authors in [112] has suggested that appropriate number and size of clusters is essential for increasing the network lifetime. If number of cluster are low then large amount of energy consumed between cluster members and cluster heads. On other side if number of clusters are high then it consumes more resources in cluster head selection. In [113] authors have proposed data aggregation mechanism that combines clustering with directed diffusion in which different categories of data are recognized and treated different during transmission.

CLUstered diffusion with dynamic data aggregation protocol (CLUDDA) [114] it diffuses the sink queries into the clustered network in which CHs are in charge of performing in-network data aggregation. Information required for collection of data are contained in queries. It allows users for the partial collection of data as per requirement. It allows data aggregation only at a selective set of CHs instead of all CHs, which reduces energy consumption. However, redundant and overlapped data collection is the major drawbacks of CLUDDA. Clustering based heuristic for maximum lifetime aggregation protocol (CMLDA) [106] [115] resolves drawbacks of CLUDDA by partitioning the network into a set of non-overlapped clusters with constant number of cluster members (CMs).

In EEHC [103] energy efficient hierarchical clustering authors have proposed a technique that increases the lifetime of the sensor network. However, in hierarchical clustering, cluster heads are overloaded and reduces its energy sooner than other nodes. Authors in [116] have proposed an algorithm to network sensors into well-defined clusters with less energy constrained gateways nodes called cluster heads and load balancing is done among these gateways to improve lifetime of system. In this paper [117] authors have proposed an enhanced PSO based clustering energy optimization in which clustering and cluster head selection is done by using PSO algorithm. The authors in [118] proposed a sensor deployment technique and optimized it using PSO but due to execution in centralized manner increases the load of the base station.

2.3 Structure Free Aggregation

A fixed network structure may result in poor aggregation efficiency for an event-based applications with dynamic source nodes because changing structure dynamically may cause high overhead.

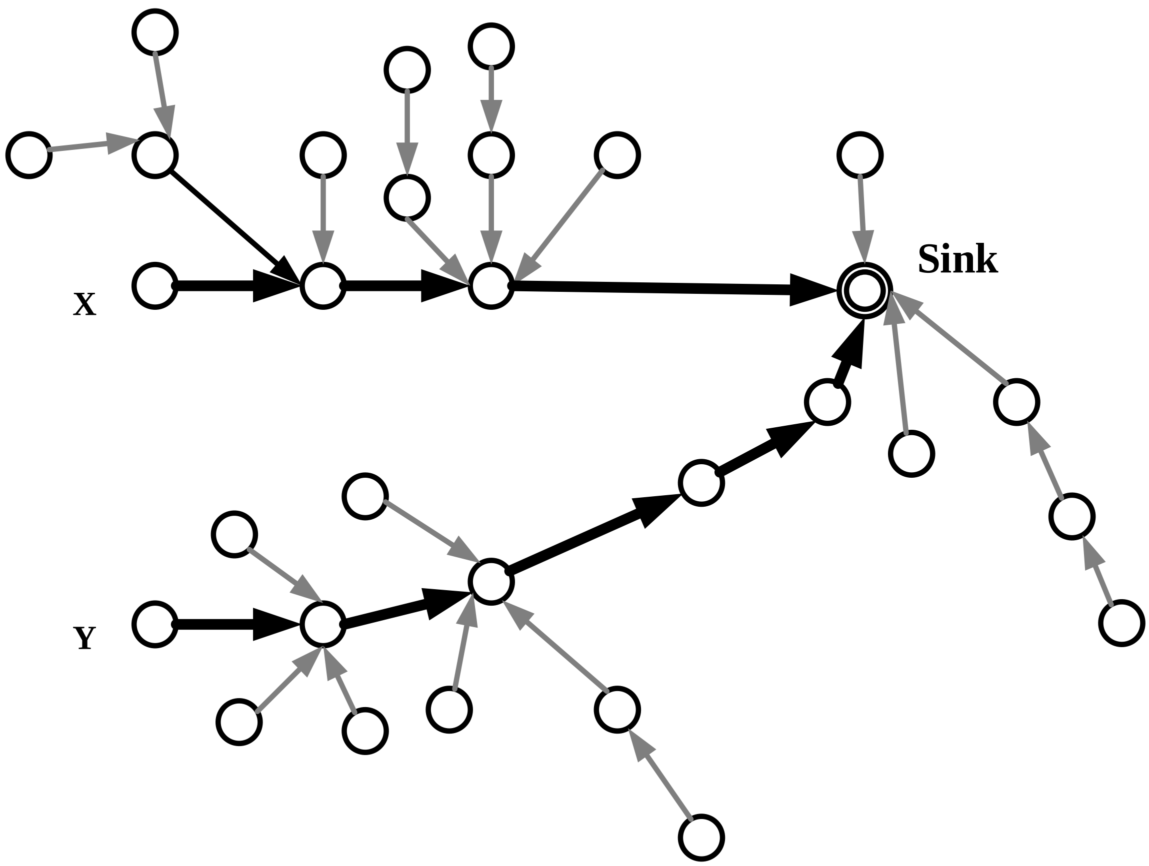

Fig. 2.2: Structure free data aggregation

For example, as shown in Fig 2.2 suppose node X and Node Y are within the range of each other their event reports take entirely different routes to the sink. If nod Y sends its event report to node X and node X aggregates received data with its own and forward the result to the sink node then only four transmissions are needed to report the event to the sink assuming that event report can be aggregated into single packet. In [73] [119]potential of structure free data aggregation explored. Here two mechanisms are proposed in structure free protocol to improve spatial and temporal convergence: Data aware anycast (DAA) and randomized waiting (RW), because there is no pre-configured structure, routing decisions are made on the fly through DAA. An RTS with an aggregation ID (AID) is broadcasted to its neighbors when a node has something to transmit. Packets with same AID can be aggregated for the improvement of spatial convergence the node's neighbors are assigned different priority to respond with CTS. Highest priority is given to neighbors having packet with same AID and closer to the sink, followed by neighbors that have packets with same AID but farther away from sink. Neighbors that have no packets with same ID and are farther away from the sink are assigned the lowest priority. Structure free aggregation even does not have any overhead in construction of structure and its maintenance but it does not guarantee aggregation of all packets from a single event, which results wastages of energy in transmitting un aggregated packets to sink.

2.4 Deployment Based Aggregation

In [120] authors have used the ABC metaheuristic to direct the movements of sensors in a sensing field for maximizing coverage area. The sensor network used in this paper comprised of both static and mobile sensors. In paper [121] authors have addressed the problem of autonomous deployment of mobile sensors that need to cover a predefined goal with a connectivity constraint. The relative neighborhood graph reduction is chosen to provide global connectivity locally. The node positioning approach based on genetic algorithm proposed in [122] . The authors used a multi-objective fitness function whose parameters included field coverage, sensor overlap error, network energy etc. Every positioning was coded as bit string sequences and each was evaluated using the fitness function.

In paper [123] a new approach was suggested by authors based on probabilistic localization technique along with virtual forces (VF) for maximizing coverage area. The VF algorithm uses a force directed approach for moving the sensors to improve the coverage. The VF algorithm had the advantage of minimal computational overhead and one-time sensor repositioning. The researchers of the paper [124] used a metric called DT-score based on it deployment sequence was generated. DT- score aims at maximizing area coverage in a static environment with obstacles. The initial deployment of sensors was done using a contour based method to minimize the number of holes, which were later filled using delaunay tri-angulation.In [125] authors have proposed a new approach that used particle swarm optimization (PSO) for maximizing coverage area at the same time decreasing the distance between the sensors. Authors of paper [126] put forward a new approach for energy efficient coverage in WSN using distributed PSO and simulated annealing (SA).The fitness of a solution was gauged on the parameters of coverage and energy consumption. For reducing the energy intake, the authors proposed the use of a hybrid algorithm comprising of PSO and SA. The local best and the global best solutions of the PSO are calibrated and corrected using SA which is per-formed on the nodes of the sensor network. In paper [127] dynamic positioning technique for coverage control in wireless sensor networks was proposed by authors based on biogeography-based optimization (BBO) meta-heuristic. In BBO, the initial solutions are called 'habitats 'and the fitness function is called the habitat suitability index (HIS).Two operations called mutation and migration were used to generate new solutions. In paper [128] authors have suggested an approach that not only guides the mobile nodes to cover the holes by calculating their best position but also ensures minimum number of mobile sensors required to achieve the objectives. The have proposed a GA based algorithm for reducing the number of holes left after the random deployment of static sensor nodes. In paper [129] authors suggested a memetic based multi-objective optimization of the coverage problem in sensor networks. In their work, the authors proposed the use of multiple local searches to find better deployment sequences that had high area coverage, efficient node utilization, and increase network lifetime. In [130] authors formulated a new approach for optimizing area coverage problem using hybrid particle swarm optimization and differential evolution algorithm. PSO is used by the hybrid algorithm for deployment of the sensor nodes to cover maximum o the preferential target points in the sensing area while differential evolution is used to carry out the cross over and mutation operations on the solution generated by the PSO algorithm. To find the better positions of some of the sensors in the deployment sequence that was generated by PSO, mutation and crossover operations used. In [131] authors proposed two algorithms for supporting distributed wireless network. Their approach deals with the problem of coverage and optimization subject to limitations such as uncertainty in detection and irregularity of terrain. The mission area divided by the authors into grids. In [132] authors proposed a computational geometry based mechanism for solving the sensor node deployment problem. In this, authors used Voronoi diagram and stationary nodes with the same sensing radius. Monitoring space is divided into cells using Voronoi diagram once stationary nodes are placed then genetic algorithm is used for placing additional mobile sensor nodes inside the cell. The fitness function used to minimize the deployments with overlaps in order to increase the area coverage. In [133] authors have suggested a new information exchange process of ABC algorithm which is being performed in one dimension with random neighbor in each solution improvement, thus convergence speed decreases as the dimension of the problem increases. In [134] authors have suggested a dynamic positioning strategy for coverage control in wireless sensor networks using biogeography-based optimization meta heuristic. Authors used the binary sensing model and a network model consisting of both static and mobile sensor nodes. In [135] authors has proposed population based optimization approach to increase coverage area of the network with electromagnetism like algorithm. In [136] authors have proposed glowworm swarm optimization based node placement algorithm with aim to maximize coverage for a fixed number of mobile nodes. This work is inspired by the behavior of glowworm that carries a luminescent substance called luciferin. The movement of glowworm based on the intensity of luciferin possessed by its neighbors. This algorithm is dealing with communication range for coverage but theoretically, in order to achieve maximum coverage the technique should be dealing with sensing range because communication range is normally associated with connectivity.

2.5 Routing Based Aggregation

Due to the vast dynamic nature of wireless channel, traditional metrics used such as hop count generally hard to provide a highly reliable path estimation in Wireless Sensor Networks (WSNs) and most probably consumed more energy in communication [137]. In paper [138] authors have proposed data routing for in-network aggregation that has some key aspect such as reduced number of messages for setting up routing tree, high aggregation rate and transmission.it also target to maximize number of overlapping routes. Proactive, reactive and hybrid are three main categories of routing, proactive algorithm Destination sequenced distance vector (DSDV) [139] [140] have a slow reaction in a network with more failure and restructuring [141]. The reactive or on-demand routing algorithm uses flooding technique for finding only one path from sender to receiver, due to flooding technique it imposes overhead on the network and high latency time in route finding. In [142] authors have proposed hybrid routing, which is a combination of reactive and proactive routing protocols. AODV (Adhoc on demand distance vector routing) is one of the most commonly used routing protocol for both MANETs and WSNs. However, in WSNs, when the mobility is high, AODV needs to find new paths to the destination frequently as the already established links breaks because of node movement [143] . Ant colony [63] is a standard solution for finding optimal path from source to destination that is proposed by authors for solving optimization problems such as travelling sales man problem (TSP) with multi-agents. In [140] authors proposed a novel routing approach using an ant colony optimization algorithm which uses artificial ants. Each ant chooses the next hop, moreover the pheromone concentration amount attends to the node's remaining energy by this method, the ant selects a node with longer lifetime. AntNet [144] is based on the principle of ant colony optimization. In AntNet, each node maintains a routing table and an additional table containing statistics about the traffic distribution over the network. In [145] the AntHocNet algorithm has been proposed which consists of both reactive as well as proactive components. In AntHocNet a source node reactively sets up a path to a destination node at the start of each communication session. HOPNET [146] is another routing protocol in mobile ad-hoc networks works by dividing the network into zones. It is a hybrid routing protocol designed to be scalable. The HOPNET algorithm consists of the local proactive route discovery within a node's neighborhood and reactive communication between the neighborhoods. The Intra-zone Routing Table is proactively maintained so that a node can obtain a path to any node within its zone quickly. This is done by periodically sending out forward ants to sample paths within its zone and determine any topology changes (such as nodes moving away, link failure, new nodes entering the zone, etc.). Once a forward ant reaches a destination, a corresponding backward ant is sent back along the path discovered. The Inter-zone Routing Table stores the path to a node not within its zone. This routing table is setup on demand when routes outside the current zone are required. The peripheral nodes of the zone are used to find routes between zones. In [147] competitive clustering technique with base station mobility was proposed to increase residual energy in sensor nodes and improve the network performance. This technique forms cluster in small near the fixed sink node that makes the head node be closer to the base station and consumes lower energy during data gathering between the clusters. To evaluate its performance, HOPNET was compared against AODV and it was found out that the HOPNET incurred higher overhead AODV for small network sizes and that the zone radius had a significant impact on packet delivery ratio, making it necessary to select this parameter carefully. Moreover, two routing tables need to be maintained for each node, hence requiring more memory. In [148] a novel Fuzzy ACO (F-ACO) routing algorithm is proposed. It extends the idea of using fuzzy logic in ant colony based protocol to present a multi-objective routing algorithm in MANETs for finding the most preferred route by evaluating the alternatives against the multiple objectives and selecting the route which best achieves the various objectives. The different routing metrics are taken into consideration using a fuzzy logic system to compute pheromone-capturing interplay of various metrics. Selection of the path performed by selecting the route having highest concentration. The inputs to the fuzzy controller for routing are buffer occupancy, remaining battery power and signal stability. A fuzzy set of rules is defined and the fuzzy cost calculated, which then decides the route cost. The inverse of the route cost gives the pheromone values to be deposited on the links. In [149] an on-demand power-balanced routing algorithm is presented for mobile, multi-hop ad-hoc networks based on the ant colony based meta-heuristic. The main goal in the design of the protocol is to reduce the overhead for routing. In [150] authors have proposed a new routing algorithm based on estimated distance to replace the hop count for the selection of the next node during the packet transmission. The use of the fuzzy logic to optimize the metric used in routing methodologies for sensor networks is a important technique since it permit us to combine nd evaluate diverse parameters in an effective manner [151]. In [152] authors have proposed a research work based on graph theory and particle swarm optimization in multi hop wireless sensor networks. The cluster head is elected with help of weighted function, which is computed iteratively for each i^{th} round. The routing of packets in the sensor networks optimized with the fitness function that target to maximize distance and residual energy after every round of transmission. In paper [153] authors have presented an overview of research based on swarm intelligence in both wired and wireless ad-hoc network. They conclude that the implementation of proposed algorithm to real-life problems is still too limited.In [64] [65] originally proposes fuzzy set theory and it has been developed for extended linguistic values. The linguistic values are terms that are used instead of the numbers and fuzzy set theory. The use of the fuzzy logic to optimize the metric used in routing approaches for wireless sensor networks is a promising technique since it allows us to combine and evaluate different parameters in an efficient manner. In [154] authors proposed a novel routing approach consisting of stable nodes based on ant colony optimization algorithm for wireless sensor networks to maximize network lifetime. Considering the energy levels of the nodes, a multipath data transfer is also accomplished to provide reliable network operations. In [155] authors have adopted conventional ant routing algorithm for wireless sensor networks by considering their traffic pattern and other constraints. In this, individual sensors are not given the task for finding its own location just to save the energy the proposed algorithm affects the task of route discovery to the base station that periodically sends forward ants over the network to discover the routes and inform the sensors about its location. In [156] authors proposed a mechanism for routing in wireless sensor network, which can be more impressive in respect of route length, end-to-end delay and network node energy. The proposed mechanism uses local inquiry and ant colony based routing algorithm to find optimal routes and to determine the route quality a fuzzy inference system was used.

Based on the analysis of the existing state-of-art algorithms for cluster-based data aggregation and fusion in wireless sensor networks, we have found the following problems and proposed the solutions for these problems:

- In literature most of the work done based on probabilistic clustering techniques but these techniques do not take into account the relative positions of selected cluster heads; and uses fixed inter cluster head election latency and also not considered differential & Temporal factors for a node and cluster, so there is a fair chance that in several rounds a considerable number of the CHs are either in proximity or very far from each other, which gives rise unbalanced clustering and wastage of energy in inter-cluster communication**,** which needs to be addressed.
- Partitioning the network into an optimal number of clusters and selecting an optimal set of nodes as CHs is N-P hard problem, which makes it a suitable candidate for the application of enhanced evolutionary algorithms to optimize energy conservation and redundancy elimination.

- A crucial issue faced in sensor resource management is that of determining the optimal positions of the sensors to maximize area coverage and it becomes harder when sensor nodes are mobile, most of the work done in the past considered static deployment, which is not suitable for mobile environment so the problem of dynamic deployment of WSNs becomes increasingly important which needs to be explored.
- In recent past most of work done on WSNs focuses on mobile ad-hoc sensor networks due to their wide range of potential applications but power consumption still remains one of the critical issues for routing in mobile ad-hoc sensor networks, which needs to be enhanced.

2.6 Research Objectives

Based upon the analysis of the earlier state-of-the-art methods of the cluster based data aggregation and fusion is global process of gathering and routing information through network with the objective of optimizing resource consumption, so we have set the following research objectives to bridge the research gap.

1) To develop a technique for the improvement of energy efficiency of probabilistic clustering technique.
2) To explore and investigate efficient collection and transmission of data with soft computing techniques.
3) To develop a model for the dynamic positioning of sensor nodes with coverage maximization using ABC algorithm.

CHAPTER 3

PROBABILISTIC CLUSTERING FOR DATA AGGREGATION

A widely used class of clustering techniques is probabilistic clustering in which a predetermined optimal probability is used to facilitate the cluster head selection process. This chapter presents a technique that improves the energy efficiency of probabilistic clustering algorithms by optimizing the number of clusters and the distribution of cluster heads in the network. Two generic approaches are also discussed to integrate proposed technique into the existing probabilistic clustering algorithms.

3.1 Introduction

In most wireless sensor network (WSN) applications, the entire network must have the ability to operate unattended in harsh environments in which pure human access and monitoring cannot be easily scheduled or efficiently managed and it's not even feasible at all [2, 157]. There is a crucial need for scalable and energy efficient routing and data gathering and aggregation protocols in corresponding large-scale environments. Sensor nodes are generally deployed randomly uncontrolled means in many significant wireless sensor networks and they form a network in an Ad-hoc manner [158]. Sensors in such networks are battery powered, energy constrained, and their batteries usually can not be recharged. Therefore we need energy aware routing and data gathering protocols that offer high scalability and low energy consumption for a long network lifetime. Sensors sense their environment and send their sensed data to a processing center, called as “Sink” or “Base Station” where all the data is collected and processed. Many routing algorithms have been proposed for efficient transmission of data between base station and sensor nodes. Grouping of sensor nodes into clusters has been widely used by researchers to satisfy the scalability, high-energy efficiency and prolong network lifetime objectives [43, 95, 110, 159]. In clustering the whole sensor network is partitioned into multiple groups of sensor nodes. Each group is called a cluster and each cluster has a leader called cluster head that perform special tasks such as data aggregation and fusion. Clustering has numerous other benefits and corresponding objectives, In addition to supporting network scalability and decreasing energy consumption through data aggregation. The route setup can be localized within the cluster and thus the size of the routing table stored at the individual sensor node can be reduced. Clustering also conserve communication bandwidth by limiting the scope of inter-cluster communication among Cluster heads (CHs) and reduces redundant exchange of information among sensor nodes. Clustered architecture is useful for sensor networks because of its inherent suitability for data fusion. The data gathered by all members of cluster can be fused at cluster head. A Cluster head can schedule activities in a cluster so that node can switch to low power sleep mode and reduce the energy consumption [157, 158].

3.2 Related Work

3.2.1 Low Energy Adaptive Clustering Hierarchy (LEACH)

LEACH is one of the most popular clustering protocols proposed for WSNs. LEACH uses rotation of cluster heads to balance network energy consumption. The operation of the LEACH is divided into a number of rounds. Each round includes a set-up phase and a steady phase. Clusters are organized in set-up phase while sensed data are transferred from sensors to cluster heads in steady phase. Fig 3.1 and fig 3.2 shows setup and steady state phase of LEACH respectively.

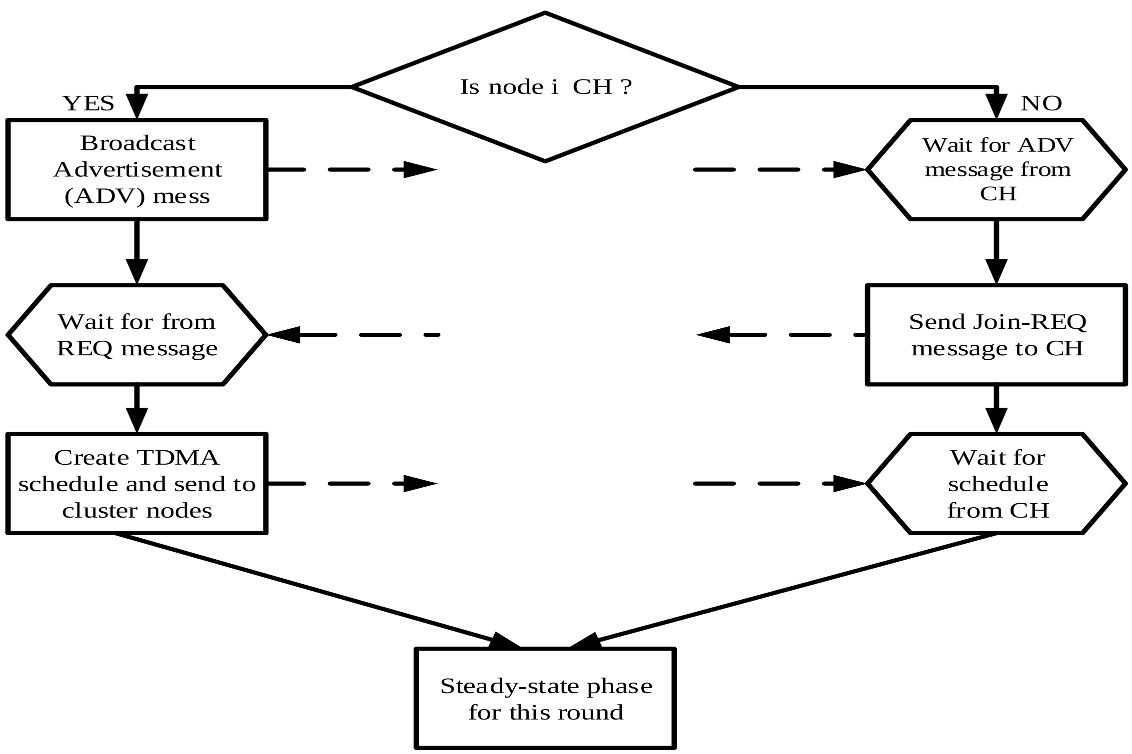

Fig. 3.1 Set up Phase of LEACH

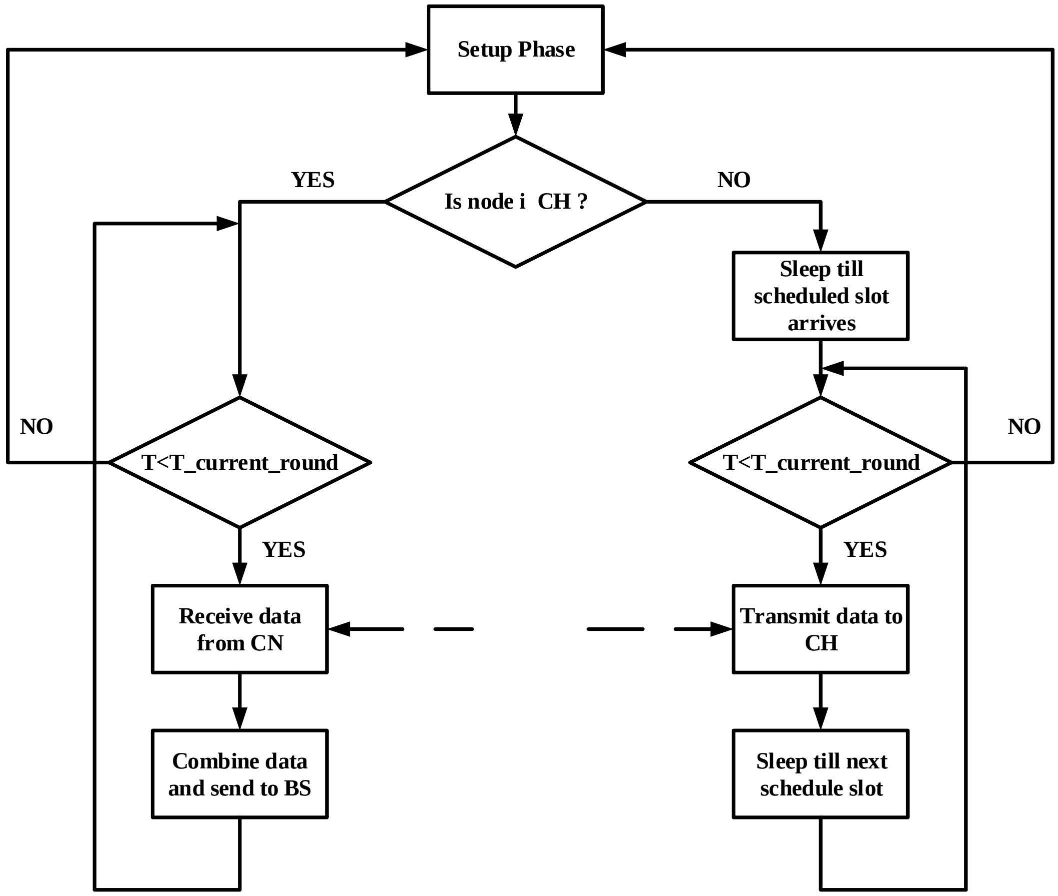

Fig. 3.2: Steady State Phase of LEACH

In LEACH, cluster heads are elected in a distributed manner and they are independent from each other. Nodes make their decision to become a cluster head by choosing a random number *r* in interval [0, 1]. A sensor node decides to be a CH in a round if its generated random number is less than the below a threshold [34]:

$$T(n) = \begin{cases} \frac{p}{1-p\left(r \bmod \frac{1}{p}\right)} & if\ n \in G \\ 0 & otherwise \end{cases} \tag{3.1}$$

Where p is the desired percentage of CH nodes in the sensor population, r is the current round number; G is the set of nodes that have not been CHs in the last 1/*p* rounds. The clusters are formed dynamically in each round and the time to perform the rounds is selected randomly. Generally, LEACH can provide a uniform load distribution in one-hop sensor networks. Moreover, it provides a good balancing of energy consumption by random rotation of CHs. LEACH uses the concept of "cluster heads" to reduce energy dissipation in information transmission, No conflict in transmission of information from cluster nodes to the cluster heads occurs due to use of TDMA, No conflict in transmission of information from cluster heads to gateway/base station node occurs due to use of CDMA, but it has some drawbacks:

- Because of probabilistic decision on cluster head election there is a good chance that a node with very low energy gets selected as a CH.
- Because of same reason, it is possible that the elected CHs may fall in one part of the network and some nodes may not have any CH in their range.
- CHs are assumed to have a long communication range so that data can reach the BS directly but this is not always a realistic assumption because the CHs are usually sensors and the BS is often not directly reachable to all nodes.
- LEACH uses a fixed inter cluster head election latency.
- LEACH does not take into account the differential energy and temporal factors for a node and cluster.
- LEACH chooses the node with the maximum remaining energy and completely ignores the communicational energy dissipation for a node.

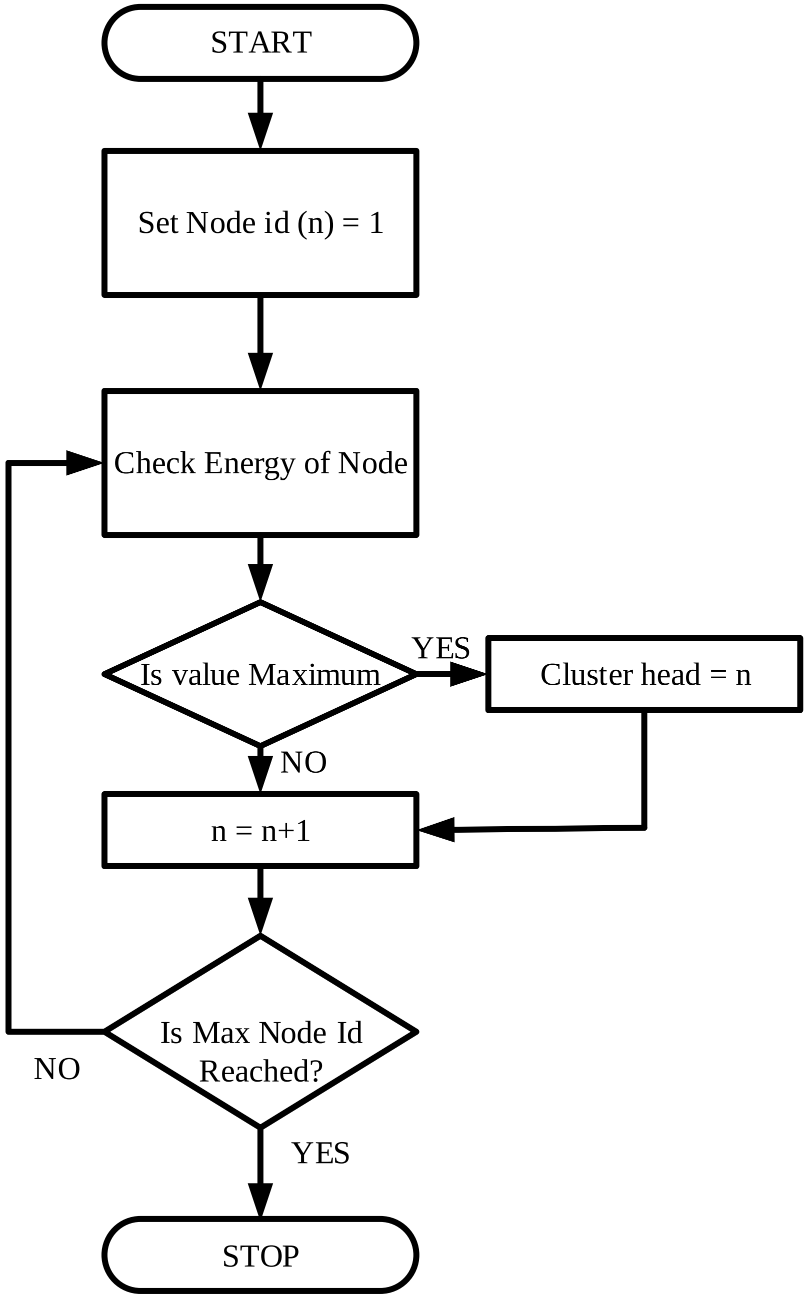

Fig. 3.3: LEACH operation

3.2.2 Hybrid Energy Efficient Distributed Clustering (HEED)

HEED is another improved and very popular energy efficient protocol. It allows multi hop communication among cluster heads (CHs) and base station (BS). HEED is a hierarchical, distributed, clustering scheme in which a one-hop communication mechanism is retained within each cluster [88]. There are two basic parameters, residual energy and intra cluster communication cost to choose CH nodes. The initial set of CHs obtained by using residual energy of nodes. On the other hand, intra-cluster communication cost reflects the node degree or node's proximity to the neighbor and is used by the nodes in deciding where to join a cluster or not. Thus, unlike LEACH, it does not select Cluster head nodes randomly. Only sensors that have a high residual energy can become CH nodes. In HEED, each node is mapped to exactly one cluster and directly communicate with its cluster head. Energy consumption is not assumed to be uniform for all the nodes. The algorithm is divided into three stages.

3.2.3 Stable Election Protocol (SEP)

SEP is a heterogeneous proactive protocol proposed in [105] with two levels of heterogeneity. It considers two types of sensor nodes, normal nodes and advance nodes. Advance nodes have more energy as compared to normal nodes. In SEP both types of nodes have weighted probabilities of becoming cluster heads. Advance nodes have more chances of becoming cluster head than normal nodes. Weighted election probabilities of the normal and advance nodes, P_{nrm} and P_{adv} can be calculated as:

$$P_{nrm} = \frac{p}{1+\alpha m} \tag{3.2}$$

$$P_{adv} = \frac{p(1+\alpha)}{1+\alpha m} \tag{3.3}$$

3.2.4 Threshold Sensitive Energy Efficient Network Protocol (TEEN)

This protocol focuses on information aggregation rather than on cluster formation. It provides a hierarchical clustered structure, grouping nearby nodes within the same cluster. TEEN [100] use data centric method with hierarchical approach. The protocol defines two thresholds: the hard threshold is a threshold (absolute) value for the sensed attribute, while the soft threshold is a threshold (small change) value of the sensed attribute. The soft threshold can be varied, depending on the criticality of the sensed attribute and the target application. The nodes transmit sensor readings only when they fall above the hard threshold and change by given amount (soft threshold). In TEEN, medium continuously sensed by sensor nodes, but data transmission is done less frequently which favours the energy saving. However, if the thresholds are not crossed, the nodes will never communicate.

3.2.5 First Order Radio Energy Model

We assess our approach using the first order radio energy model [34]. The Fig. 3.4 shows that $E_{Tx}(m, d)$ is the energy spent to transmit an m-bit packet between a transmitter and receiver along the distance, d is

$$E_{Tx}(m, d) = \begin{cases} m * E_{elec} + m * E_{fs} * d^2 \ : \text{ if } d < d_o \\ m * E_{elec} + m * E_{mp} * d^4 : \text{ if } d \geq \ d_o \end{cases} \tag{3.4}$$

Where $d_o = \sqrt{E_{fs}/E_{mp}}$ signifies the threshold distance, E_{elec} signifies the energy loss for forwarding or receiving one bit, $E_{fs} * d^2$and $E_{mp} * d^4$ means power loss for free space and multipath fading channel.

$E_{Rx}(m)$ is the energy wasted by the receiver to obtain an m-bit message is given by

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

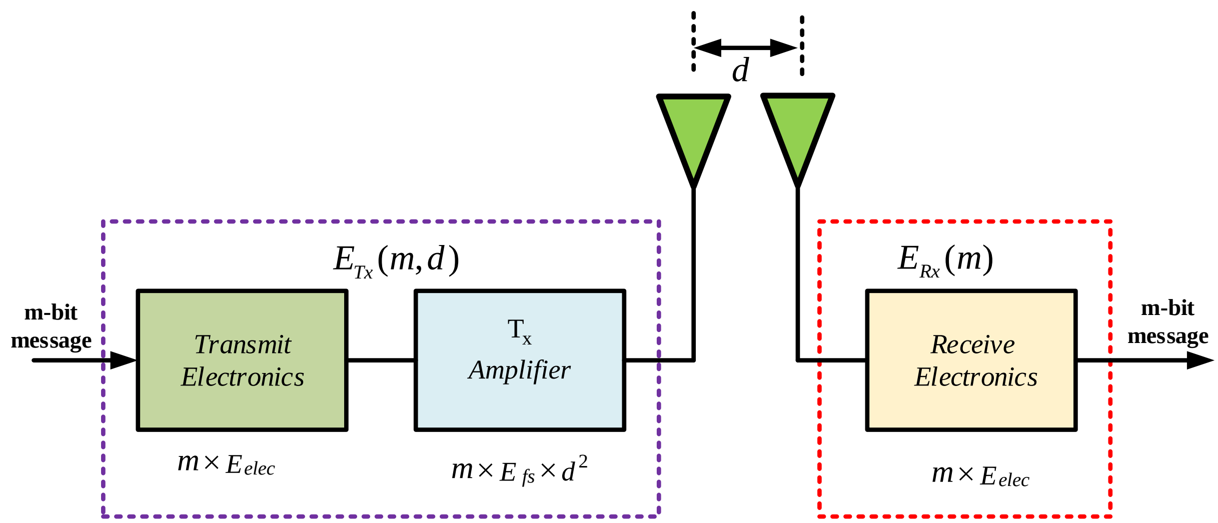

Fig. 3.4: First Order Radio Energy Model

3.3 Proposed Work

In this section, the detailed reasoning of proposed Energy Efficient Probabilistic Clustering Technique (EEPCT) for Data Aggregation in Wireless Sensor Network is discussed. We have used the same radio energy model for power consumption in radio transmission as used in most of the probabilistic Clustering algorithms such as LEACH [34].

3.3.1 Requirement of Clusters of Almost same size

Probabilistic clustering techniques usually do not take into account the relative positions of selected cluster heads; as a result, there is a fair chance that in several rounds a considerable number of the CHs will be either in proximity or very far from each other. In case, when a considerable number of cluster heads are in close proximity, the number of nodes in the clusters associated with these cluster heads are considerably low in comparison to other clusters, which give rise to uneven clustering. While a well-balanced clustering i.e. nearly same sized clusters is crucial for reasonable performance of probabilistic clustering techniques such as LEACH, TEEN etc. On the contrary, if selected cluster heads are very far from each other, there will be a significant wastage of energy in inter-cluster communication in protocols such as HEED, EEHC etc.

Therefore, a uniform distribution of cluster heads over entire network region is desirable to achieve a well-balanced clustering.

Therefore, as our first proposed improvement, we want to introduce a parameter ξ to denote the closeness and its value depends on the size, node density and number of cluster heads to be selected. If in a particular round the distance between any two selected cluster heads is smaller than ξ, they are considered too close to each other and one of them has to drop its decision of becoming cluster head (CH).

3.3.2 Maintaining the Optimal Percentage of CHs in Each Round

Though the above proposed improvement in cluster head selection process causes a significant improvement in the performance of probabilistic clustering algorithms, it has a small drawback. Consider the case when a large number of potential CHs (the nodes for which the value of generated random number is less than the threshold value) in a round of a particular epoch are in close proximity then many of them will drop their decision to become CH. In such a situation, the number of selected cluster heads would be significantly lesser than the optimal number of cluster heads. This will lead to bigger size clusters and consequently more energy consumption in intra-cluster communication. Also in the last round of the epoch all the nodes that have not become cluster head so far in that particular epoch will have to become cluster head and so chances will be higher for the selection of more cluster heads than optimal number of cluster heads. This will increase the long distance transmission to the sink. To deal with this situation we want to increase the number of the potential cluster heads in each round so that the number of selected CHs after dropping some of potential CHs because of ξ-closeness be as near to optimal value as possible. For this, we need to raise the threshold value in equation 3.1 so that more nodes than usual will be eligible to become cluster head.

3.3.3 Choosing Closeness Factor ξ

As a first improvement, we have introduced a parameter ξ to denote the closeness of two cluster heads and its value will depend on size, node density of network and number of cluster heads to be selected. If in a particular round the distance between any two selected cluster heads is smaller than ξ, then they are considered too close to each other and one of them has to drop its decision of becoming cluster head. This way we can ensure that cluster heads will not be very close to each other and will be well distributed across the network.

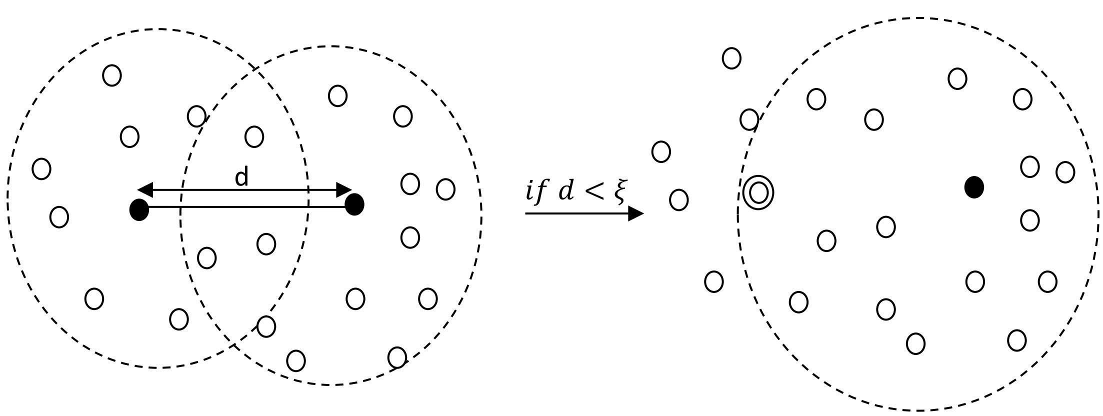

Fig. 3.5: Effect of ξ-Closeness on Cluster

3.3.4 Threshold Increment Factor σ

In our second proposed improvement, another parameter σ introduced, denoting the threshold increment factor. The increment in the threshold value given by equation 3.1 will not be same for each round in an epoch as original threshold value itself increases in each subsequent round of an epoch and becomes equal to 1 in last round but we want to keep σ constant in each round so we can calculate new increased threshold using equation 3.6

$$T_{NEW}(n) = T(n) + \sigma(1 - T(n)) \tag{3.6}$$

In our assumption, σ will be constant for a particular configuration of network and a particular choice of closeness factor ξ.

3.3.5 Determination of Closeness Factor ξ and Threshold Increment Factor σ

The values of ξ and σ are very crucial for the better performance of clustering algorithms. Closeness Factor ξ depends on the network configuration, i.e. size of the network, number of nodes deployed in the network, density of nodes, and optimal number of cluster heads, while the value of σ depends on the value of ξ as well as on network configuration. We have calculated the optimal values of ξ and σ experimentally by running multiple instances of the algorithm for different combinations of ξ and σ, keeping other parameters of the network constant.

3.3.6 Integration with the Existing Probabilistic Techniques

The above two proposed improvements integrated with existing probabilistic clustering techniques to improve their energy efficiency. We have used two generic approaches to integration, one that requires little intervention of centralized authority such as base station and other follows a completely a distributed approach.

A. **Centralized Approach***:* In Fig. 3.6 functioning of centralized approach of integration discussed.

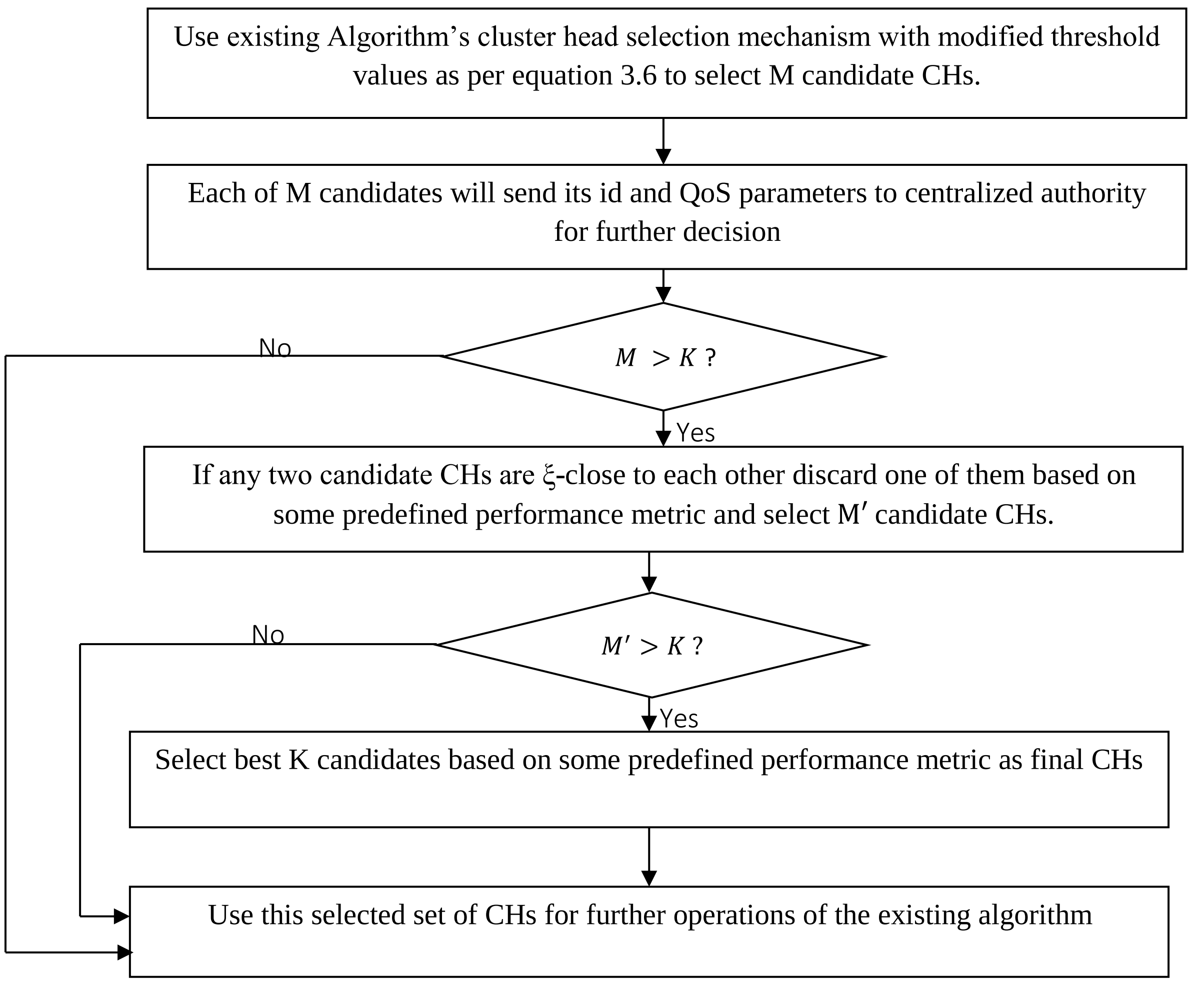

Fig. 3.6: A centralized approach to integrating proposed improvements

Here K is the optimal number of CHs that is precalculated using analytical methods as described in [34]. The predefined performance metric can include QoS factors such as quality of link between candidate and base station, congestion, node density around the candidate node, its residual energy etc.

B. Distributed Approach: in distributed approach, each node autonomously takes decision of becoming a cluster head or not without intervention of any central authority. The following Fig. 3.7 explains a distributed approach for integration.

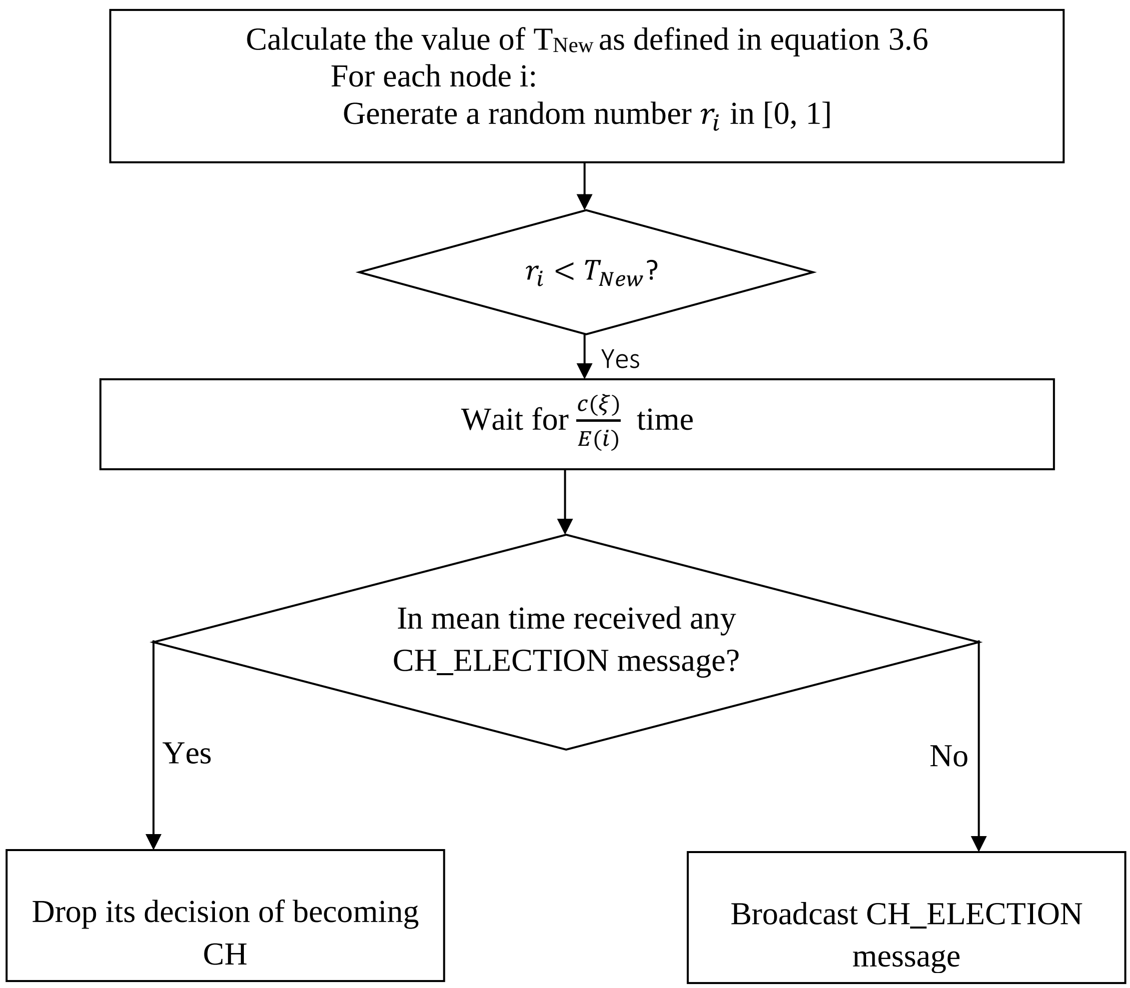

Fig. 3.7: A distributed approach for integrating proposed improvements

Here E(i) is the residual energy of node i and $c(\xi)$ is a function whose value depends only on closeness factor ξ. Waiting for $\frac{c(\xi)}{E(i)}$ time ensures that among the entire candidate CHs that are in ξ-closeness range of each other only the node with highest residual energy will be chosen as cluster head (CH).

Both centralized and distributed approaches have their own advantages. Centralized approach offers better QoS implementation mechanism and more tolerant to variance in values of σ and ξ from optimal value. On the other hand, distributed approach requires minimal message exchange among nodes for CH election. Here we consider c as a function of ξ only but for QoS implementation, we can make c as a function of other QoS parameters with ξ as well.

3.3.7 Optimal Differential Inter Cluster Head Election Latency Determination

- Conventional LEACH takes a fixed Inter Cluster Head Election Latency value.
- The proposed technique rather than taking a fixed inter cluster head election latency, considers cluster activity also to determine the same.
- Cluster Activity (A) is measured using the average value of communication latency in the cluster.
- Inter cluster head election latency(C) deduced using following equation

$$C = n * \left(\frac{T}{A} \right) \tag{3.7}$$

Here,

C=Inter cluster head election latency

n=constant dependent upon minimum possible communication latency (A_m) for the network

T=constant Inter cluster head election latency used for conventional LEACH

A=Cluster Activity

3.3.8 Optimal Differential Energy-Based Cluster Head selection

- Conventional LEACH selects a cluster head solely on the basis remaining energy of the nodes
- The proposed technique in addition to energy, also takes into account the communicational energy dissipation measured using the Cluster activity (A), as described above.
- We assign a metric value to each node, the node with the maximum value of the metric is elected as the cluster head.

$$Metric = (E)^{a}\left(\frac{1}{A}\right)^{1-a} \quad (3.8)$$

Here,

E=Energy of the node

A=cluster activity

a,1-a=weights assigned to the two components

- The value of a was calculated using ant colony successive magnification optimization algorithm.

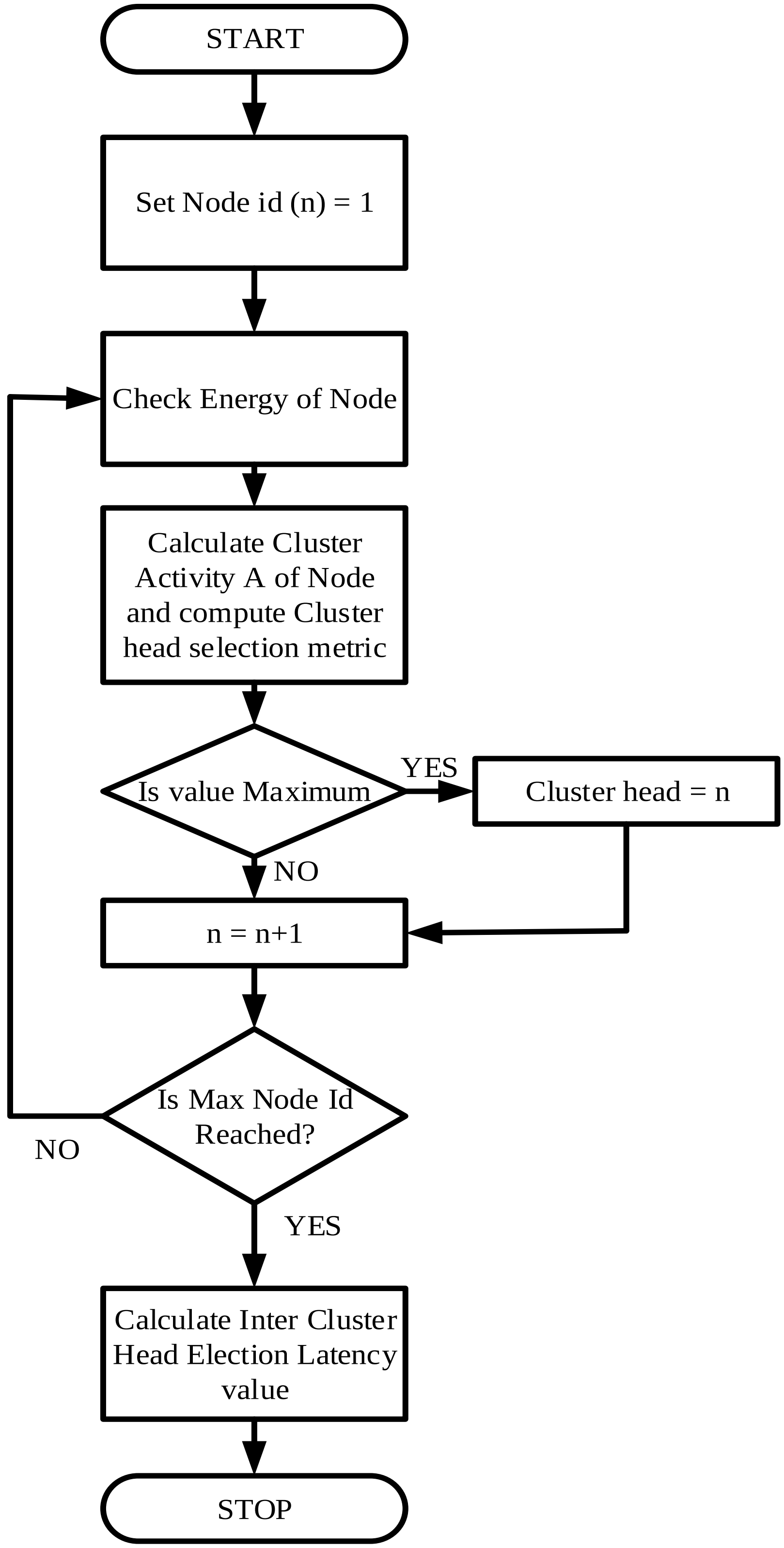

Fig. 3.8: Shows operation of Modified LEACH

3.4 Simulation Results

We conducted a rigorous simulation study to evaluate the performance of proposed Energy Efficient Probabilistic Clustering Technique (EEPCT). We assume a square network field of dimension 100m X 100m with 100 sensors deployed in it and MATLAB is used for the simulation. Specific parameters for simulation are shown in Table 3.1. The performance of EEPCT evaluated and compared with some representative probabilistic clustering algorithms such as LEACH [34], HEED [88], SEP [105].

Table 3.1: **Simulation parameters**

Description	Parameter	Value
Initial energy	E_0	0.5J
Electronic circuitry energy	E_{elec}	50nJ/bit
Multi-path co-efficient	ε_{mp}	10 pJ/bit/m2
Free space co-efficient	ε_{fs}	0.0013 pJ/bit/m4
Data aggregation energy	E_{DA}	5 nJ/bit/signal
Data packet size	P_{pkt}	30 bytes
Total no. of nodes	N	100
Optimal percentage of CHs	P_{opt}	0.1

3.4.1 Determination of ξ and σ

To determine optimal values of ξ and σ we run our algorithm multiple times for different combinations of ξ and σ for a given configuration of the network and choose best values.

Table 3.2 shows energy consumption for different combinations of ξ and σ for two different network configurations.

Table 3.2: **Values of ξ and σ**

ξ	**σ**	**Energy Consumption in Configuration 1 (Joule)**	**Energy Consumption in Configuration 2 (Joule)**
0	0.00	14.69	119.65
5	0.05	14.56	117.74
5	0.09	14.66	117.80
5	0.17	14.27	119.46
10	0.05	13.62	119.43
10	0.15	13.73	116.77
10	0.20	13.91	115.99
11	0.05	13.54	115.54
11	0.15	13.46	115.54
11	0.19	13.27	115.54
12	0.05	13.22	115.59
12	0.15	13.47	114.91
12	0.20	13.67	115.67
13	0.05	13.22	115.27
13	0.15	13.08	115.13
13	0.20	13.17	115.17
14	0.05	13.10	115.76
14	0.15	13.01	115. 58
14	0.20	12.93	115.58
15	0.00	13.23	116.44
15	0.05	13.01	116.01
15	0.10	12.66	116.62

15	0.11	12.61	116.49
15	0.13	12.63	116.49
15	0.15	12.54	116.32
15	0.16	12.51	117.51
15	0.17	12.84	117.96
15	0.19	12.73	116.59
16	0.05	12.87	116.87
16	0.15	12.69	116.76
16	0.20	12.67	116.14
17	0.05	12.95	116.97
17	0.15	12.83	116.52
17	0.25	12.83	117.01
18	0.05	12.87	116.66
18	0.10	12.67	117.67
18	0.20	12.68	116.19
19	0.05	12.93	114.93
19	0.15	13.11	114.27
19	0.20	13.09	114.81
20	0.05	13.83	115.36
20	0.20	13.59	116.94
20	0.25	13.77	116.27
22	0.05	13.69	115.07
22	0.10	13.62	115.21
22	0.15	13.71	114.71
25	0.05	14.20	112.52
25	0.15	14.25	112.65
25	0.20	14.19	111.19
27	0.05	14.30	111.69
27	0.15	14.67	112.74
27	0.25	14.83	111.73
29	0.02	14.55	110.88

29	0.04	14.07	110.86
29	0.05	14.34	110.47
29	0.07	14.39	110.39
29	0.09	14.94	110.43
29	0.10	14.99	110.59
29	0.13	14.21	110.04
29	0.14	14.69	110.54
29	0.15	15.12	110.87
29	0.16	14.99	110.51
29	0.20	14.45	112.68
29	0.21	14.56	112.37
29	0.25	14.74	114.90
30	0.05	14.69	114.12
30	0.10	15.67	115.07
30	0.20	15.19	115.43

We have calculated the optimal value of ξ and σ for two Network configurations. In first configuration, 30 sensors deployed in a 50m X 50m network field. In second configuration, we have taken a 100m X 100m network field with 100 sensors deployed in it. We found 15m and 0.16 to be optimal values for ξ and σ respectively for first configuration. Similarly, for second configuration optimal values are 29m and 0.13 respectively.

3.4.2 Evaluation of 'a'

- Since the parameter a for differential energy based cluster head selection can assume values in range [0, 1], we run the optimization algorithm over the range. The results obtained as follow:

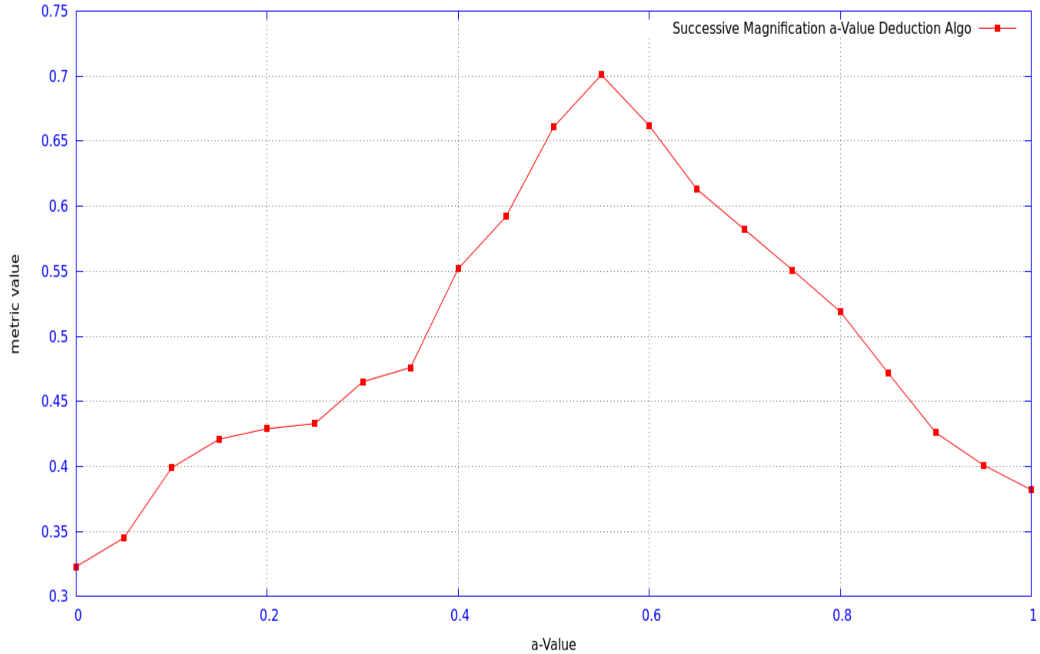

Fig. 3.9: Metric value vs a-value

Seeing that the peak occurs at 0.55 a-Value we now use ant colony successive magnification algorithm between 0.5 and 0.6.

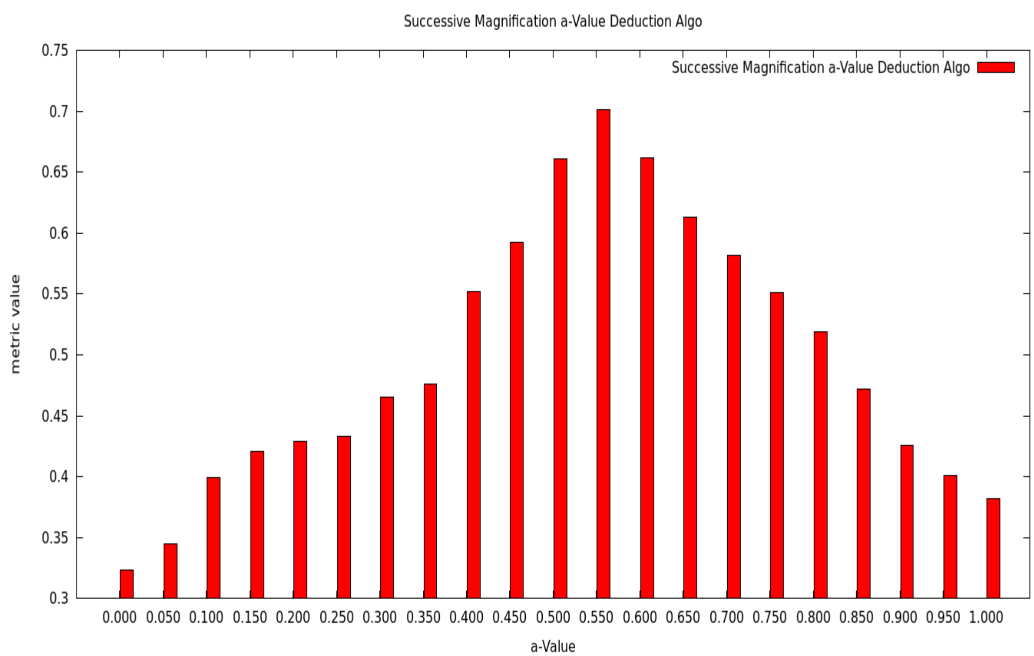

Fig. 3.10: Metric value vs a-value

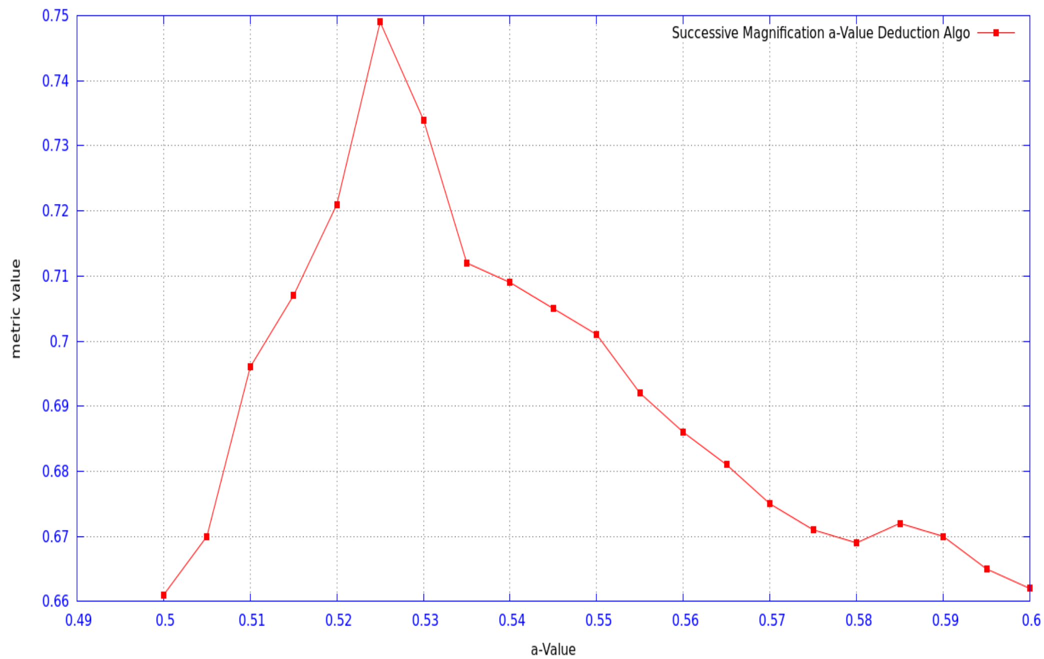

Fig. 3.11: **Metric value vs a-value between 0.5 and 0.6**

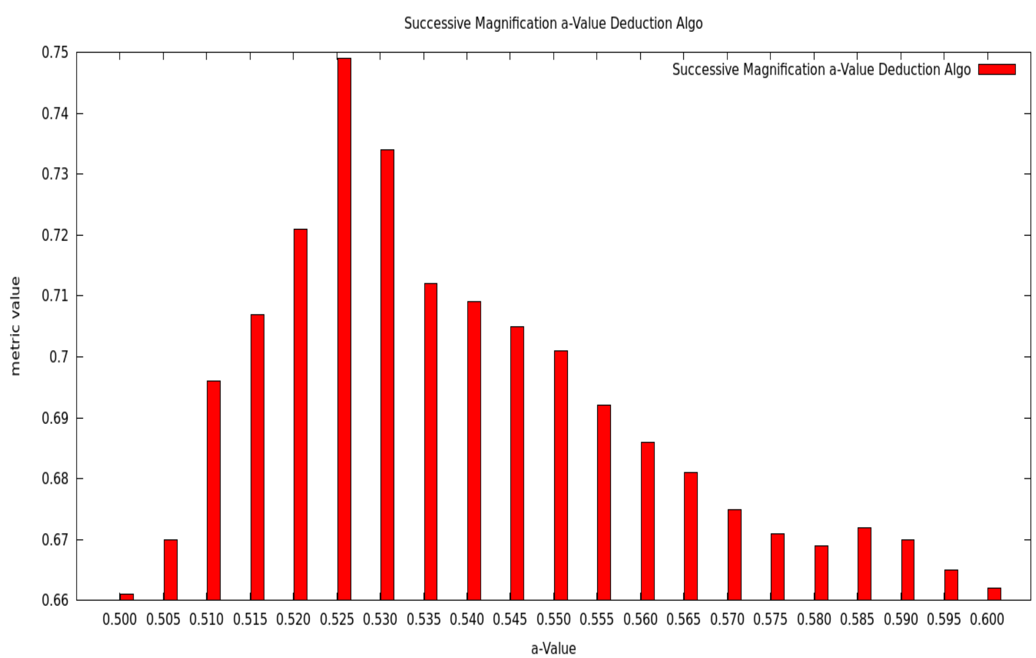

Fig. 3.12: **Metric value vs a-value between 0.5 and 0.6**

Now the acquired data is sufficient for ant colony protocol for tracing a smooth curve

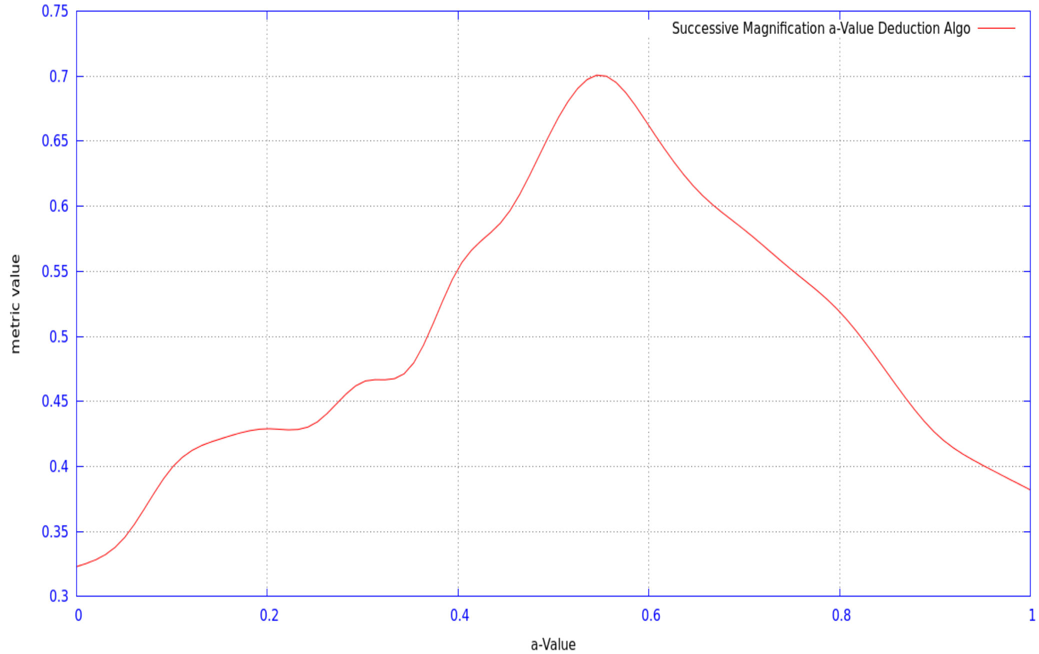

Fig. 3.13: Smooth curve between metric value and a-value

The curve has a max value at 0.525 a-Value

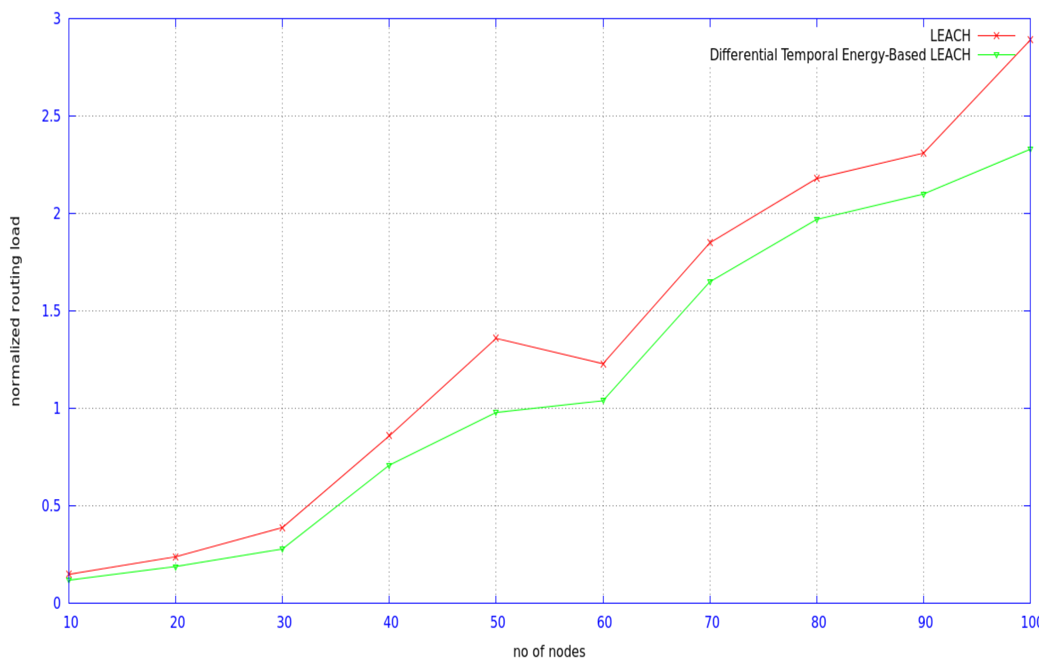

Fig. 3.14: Normalised routing load vs number of nodes

Normalised Routing Load refers to ratio of routing packets and data packets. In our modified algorithm due to the selection of an optimal zone for a node, routing packets per data packet decreases. Absolute temporal and energy analysis restrains the flexibility and efficiency of the network, which increases routing loads. Therefore, by using Differential Temporal and Spatial analysis using an optimal a-Value and appropriately calculated Cluster Activity value (A), we have achieved reduction in normalized routing Load.

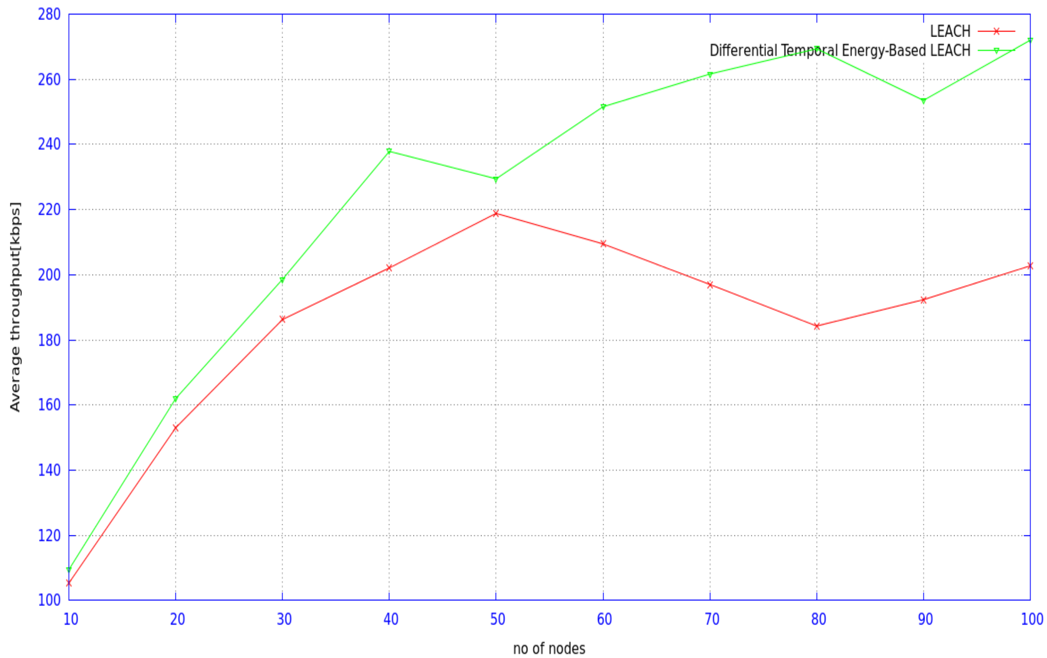

Fig. 3.15: **Throughput vs number of nodes**

Average throughput refers to the data packets delivered per unit time. Due to reduction in energy consumption by choosing a cluster head having minimum Cluster Activity (A)(by using Differential Energy-Based protocol) which leads to selection of shorter paths, delivery of packets increases with respect to time in our algorithm. This results into a better Average Throughput with respect to LEACH.

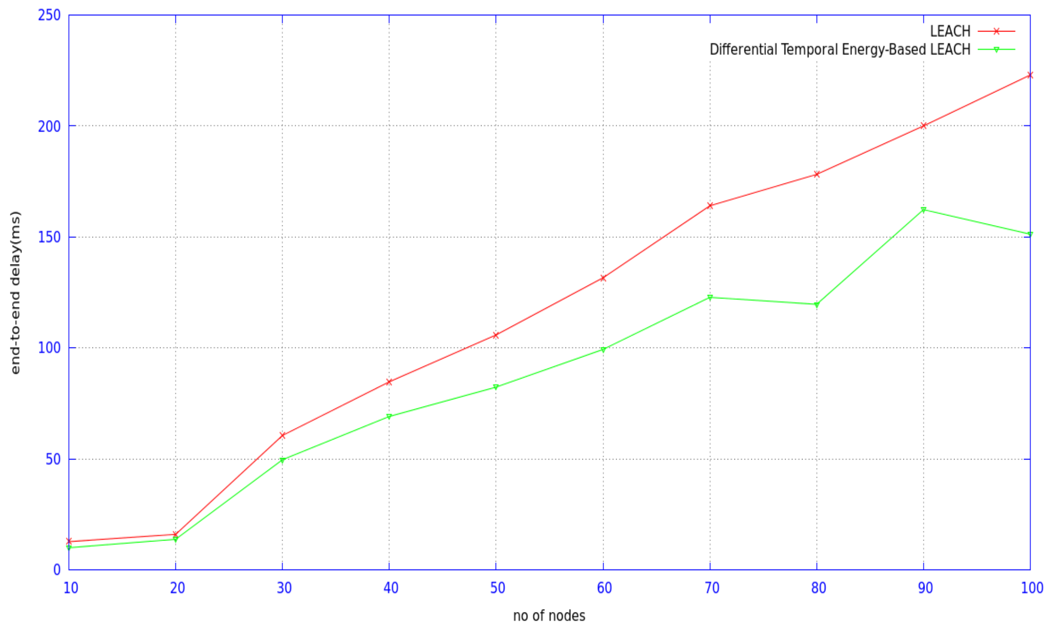

Fig. 3.16: **End to end delay vs number of nodes**

It refers to the difference in sending and receiving time.Due to election of a better cluster head following Differential Energy-Based analysis we generally choose a shorter transmission path. This results into decrease in end-to-end delay.

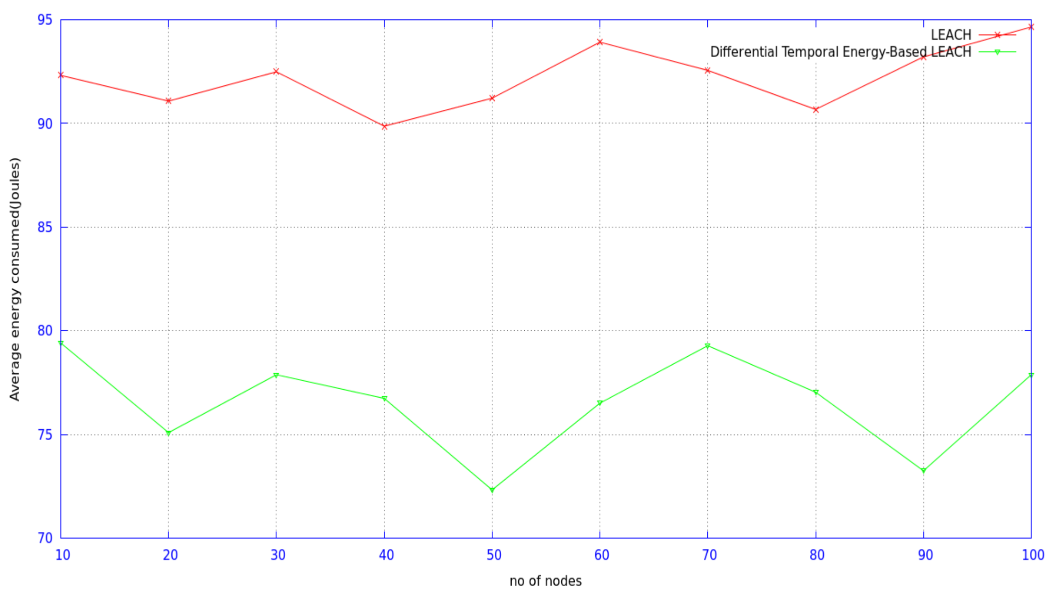

Fig. 3.17: **Average Energy Consumption vs number of nodes**

Average Energy Consumption, as the name implies, refers to the average amount of energy consumed during the WSN operation. This is probably the most important metric for comparison. The proposed protocol used Cluster Activity (A) to determine optimal inter cluster head election and optimal cluster head selection using an optimal a-Value. This results in a significant decrease in average Energy consumption.

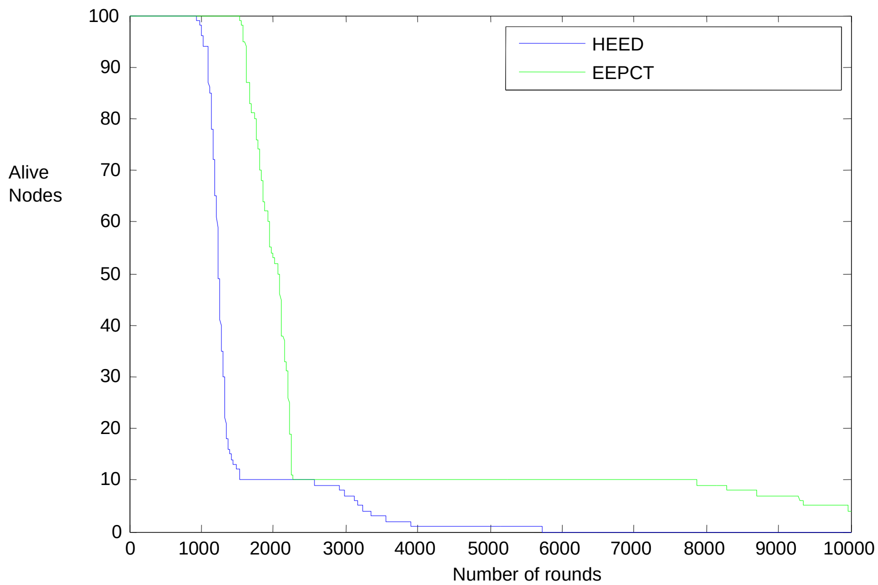

Fig. 3.18: **Number of alive nodes per round**

In above figures, we have shown the performance of existing probabilistic clustering algorithms in their original form and after the application of our proposed technique. Fig. 3.18 and Fig. 3.19 shows that after applying our proposed technique there is an increase of 43% in stability period of HEED and an overall 66 % increment in total lifetime of network in simulated environment.

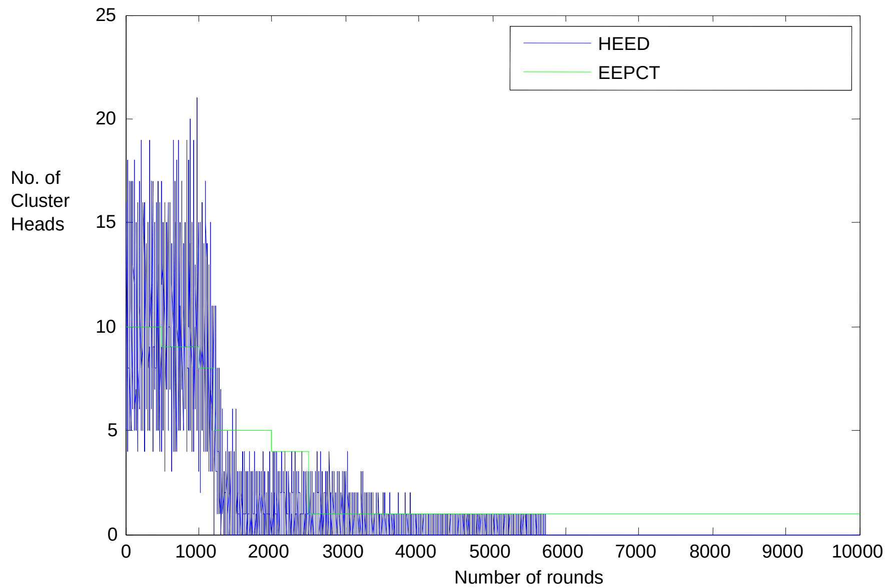

Fig. 3.19: Number of Cluster heads (CHs) per round

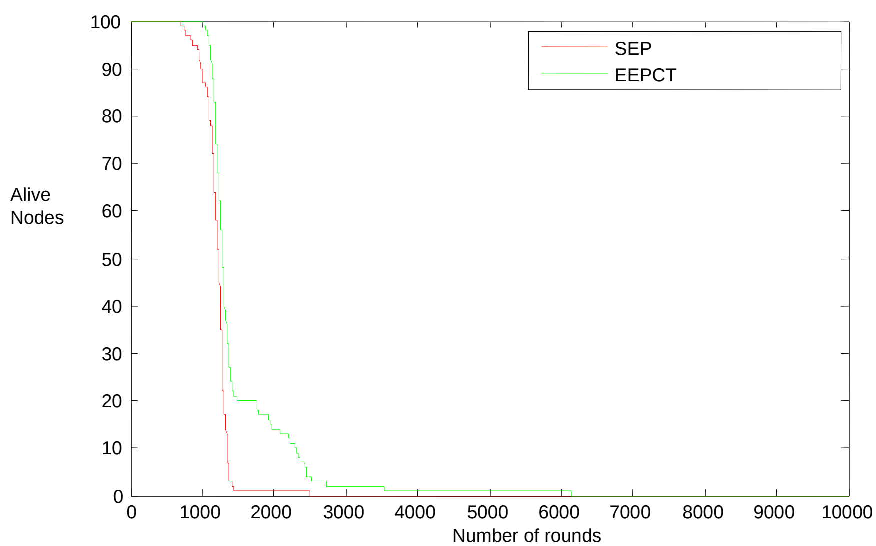

Fig. 3.20: Number of alive nodes per round

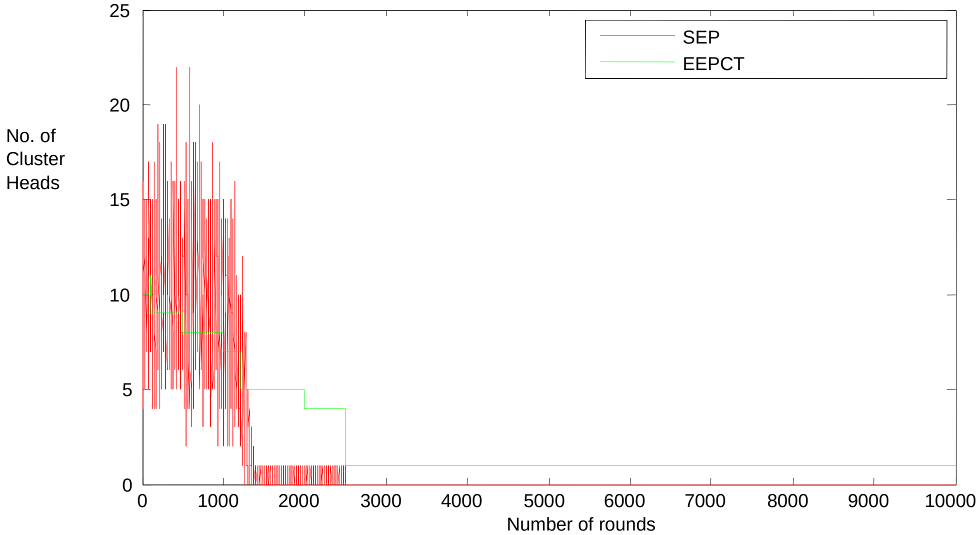

Fig. 3.21: Number of Cluster heads (CHs) per round

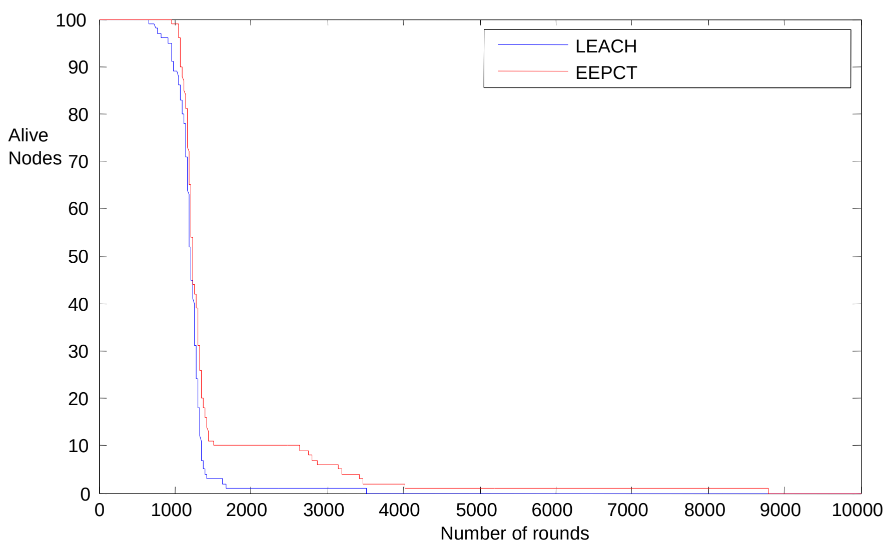

Fig. 3.22: Number of alive nodes per round

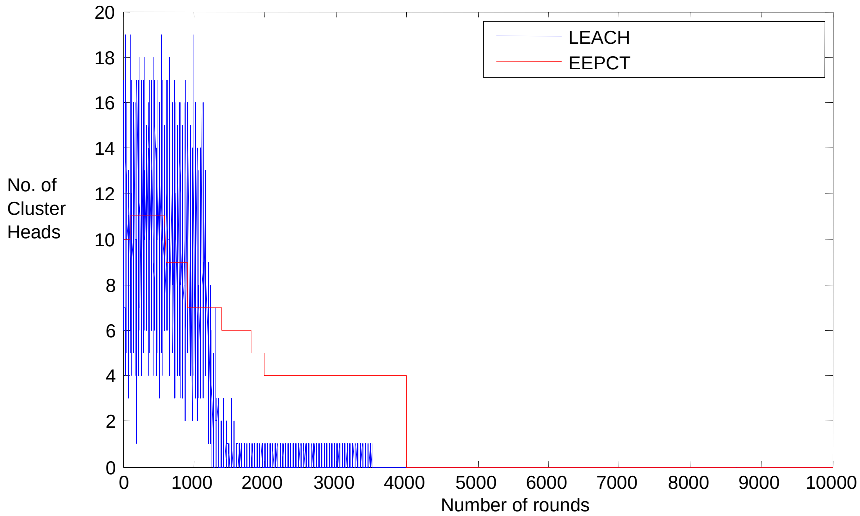

Fig. 3.23: Number of Cluster heads (CHs) per round

Similarly, in case of SEP, in Fig. 3.20 and 3.21 we can see a considerable increase in the energy efficiency and network lifetime. We can also observe that after application of our proposed modification there is a significant improvement in the no. of clusters formed in each round.

Similarly, in case of LEACH, Fig. 3.22 and 3.23 shows a significant improvement in overall lifetime of the network. There is also an increase of around 37 % in stability period of the network.

3.5 Conclusion

The proposed probabilistic clustering technique for data aggregation outperforms in comparison of state of art techniques HEED, SEP and LEACH. By using differential temporal and spatial analysis using an optimal a-Value and appropriately calculated Cluster activity value, we have achieved reduction in normalized routing Load, improvement in average throughput with respect to LEACH and significant decrease in average energy consumption. After applying proposed technique there is an increase of 43% in stability period of HEED and an overall 66 % increment in total lifetime of network in simulated environment. Similarly, in case of LEACH, there is a significant improvement in overall lifetime of the network. There is also an increase of around 37 % in stability period of the network. While in the case of SEP, significant improvement is there in network lifetime.

Published work related to this chapter:

Rajesh K. Yadav, Daya Gupta, D. K. Lobiyal, "Energy Efficient Probabilistic Clustering Technique for Data Aggregation in Wireless Sensor Network," ***Wireless Personal Communications***,2017, DOI: 10.1007/s11277-017-4370-5(Pub.: Springer), [SCI Indexed].

CHAPTER 4

CLUSTER BASED DATA AGGREGATION USING PSO

In this chapter, particle swarm optimization based clustering with discrete search space is presented in which clustering and cluster head selection are done using particle swarm optimization(PSO) algorithm, then an energy efficient hybrid clustering protocol (EEHCP) for multilevel heterogeneous wireless sensor networks are explored to minimize the power consumption in WSN.

4.1 Introduction

The main goal of data aggregation algorithms is to gather and aggregate data in an efficient manner so that lifetime of the network increases by decreasing the number of packets to be sent to sink or base station, in turn reduces the communication costs and energy consumption. Grouping of sensor nodes into clusters has been widely used by researchers to satisfy the scalability, high-energy efficiency and prolong network lifetime objectives [6, 7]. In clustering, the whole sensor network is partitioned into multiple groups of sensor nodes. Each group is called a cluster and each cluster has a leader called cluster head that perform special tasks such as data aggregation and fusion.

Particle Swarm Optimization (PSO) is an optimization technique in which natural species social behaviors are considered for the purpose of computation [56]. It is a swarm intelligence technique, which is based on population that performs optimization process with the objective of optimizing a fitness function. This approach makes use of a swarm for the purpose of search on every particle and records the fitness value of each particle. Then the particles are linked with their matching velocity. It helps the particle to make a move to a proper location by considering the optimized fitness function's cost [29, 160].

4.2 Related Work

Particle swarm optimization (PSO) inspired by social behavior of bird flocking or fish schooling is a population based stochastic technique to solve continuous and discreet optimization problems. PSO learned from the scenario and used it to solve the optimization problems. In PSO, every single solution is named as "bird" in the search space. We call it "particle". All of particles have fitness values, which are evaluated by the fitness, function to be optimized, and have velocities, which direct the flying of the particles. The particles fly through the problem space by following the current optimum particles [58, 57].

PSO is initialized with a group of random particles (solutions) and then searches for optima by updating generations. In every iteration, each particle is updated by following two "best" values. The first one is the best solution (fitness) it has achieved so far. (The fitness value is also stored.) This value is called $pbest$. Another "best" value that is tracked by the particle swarm optimizer is the best value, obtained so far by any particle in the population. This best value is a global best and called $gbest$. When a particle takes part of the population as its topological neighbours, the best value is a local best and is called $lbest$ [161] Suppose, there is a group of K random particles in an n-dimension searching space, the position of the i^{th} particle is $X_i = (x_{i_1},\ x_{i_2}, \dots \dots, x_{i_n})$, the personal best value of the particle is $pbest_i = (p_{i_1}, p_{i_2}, \ldots \ldots, p_{i_n})$, and the velocity of the particle is $V_i = (v_{i_1}, v_{i_2}, \ldots \ldots, v_{i_n})$. The best value obtained so far by any particle in the population is $gbest = (g_1, g_2, \ldots \ldots, g_n)$. After finding the two best values, $pbest$ and $gbest$ the particle updates its velocity and positions as follows

$$v_{i_j} = w.v_{i_j} + c_1.r_1\left(p_{i_j} - x_{i_j}\right) + c_2.r_2\left(g_j - x_{i_j}\right) \quad (4.1)$$

$$x_{i_j} = x_{i_j} + v_{i_j} \quad (4.2)$$

Where w is inertia and used to control the trade-off between the global and the local exploration ability of the swarm, c_1 and c_2 are learning factors, r_1 and r_2 random numbers between 0 and 1.

4.3 Proposed Work

4.3.1 Fitness Function

Success of proposed algorithm depends greatly on the formulation of fitness function. Therefore, we are defining a fitness function that includes all optimization criteria. The aim is to minimize the intra-cluster communication energy and energy loss due to cluster head and base station communication, so we can define the fitness of a particle i as

$$F(P_i) = E_1(P_i) + \mu\, E_2(P_i) \quad (4.3)$$

$$E_1(P_i) = \sum_{k=1}^{K} \sum_{\forall n_{k_j} \in C_k} \frac{f\left(n_{k_j}, CH_k\right) - E_{min}}{E_{max} - E_{min}} \quad (4.4)$$

$$E_2(P_i) = \sum_{k=1}^{K} \frac{g(CH_k, BS) - E_{min}}{E_{max} - E_{min}} \tag{4.5}$$

$$f\left(n_{k_j}, CH_k\right) = \begin{cases} s^2\left(n_{k_j}, CH_k\right) & if\ s\left(n_{k_j}, CH_k\right) \le d_0 \\ s^4\left(n_{k_j}, CH_k\right) & if\ s\left(n_{k_j}, CH_k\right) > d_0 \end{cases} \tag{4.6}$$

$$g(CH_k, BS) = \begin{cases} d_{CH_k,BS}^2 & if\ d_{CH_k,BS} \le d_0 \\ d_{CH_k,BS}^4 & if\ d_{CH_k,BS} > d_0 \end{cases} \tag{4.7}$$

$$s(n_i, CH_k) = \begin{matrix} \min(s_{n_i,CH_k}) \\ \forall k = 1,2,\dots,K \end{matrix} \tag{4.8}$$

Where, $d_{i,j}$ is the distance between node i and node j; s is a function that finds the minimum distance cluster head for a given node; f is a function whose value for a given node is proportional to the energy consumed in communication between the node and its cluster head; similarly g signifies the energy loss due to cluster head and base station communication; E_{max} and E_{min} are the maximum and minimum energy loss in the network. C_k is k^{th} cluster in a solution or particle.

E_1 and E_2 are two normalized functions that represent the energy dissipated in intra-cluster communication and due to communication between sink and CHs respectively. F is fitness function and our aim is to minimize this function.

μ is a controlling parameter for controlling the distance between base station and cluster heads. The higher the value of μ shows the closer CHs from BS. K is the optimal number of cluster heads. For each particle or solution k random nodes are chosen as cluster heads and remaining nodes joined the cluster whose CH is at minimum distance from it. Then the value of fitness function evaluated for each particle and values of $pbest$ and $gbest$

calculated. Then the velocity vector and position vector updated according to equation 4.1 and 4.2.

4.3.2 A new operator $\oplus_{NW}$

We have defined a new operator $\oplus_{NW}$ that when applied on a location with respect to a network, it returns a valid sensor node location in the network. In each iteration of our algorithm location of CHs updated in each particle or solution. Keeping this into consideration operator $\oplus_{NW}$ is defined as follows: Suppose $\dot{a} = (a_1, a_2)$ is any location with respect to a sensor network NW then $\oplus_{NW}\ \dot{a}$ returns a valid location in network NW. The operator $\oplus_{NW}$ first enquires if $\dot{a}$ is a valid location then it return $\dot{a}$ as it is; if not then it returns nearest valid location in the network NW toward base station with highest residual energy. After calculating new velocity and position, using equation 4.1 and 4.2 operator is applied to the calculated positions to get new valid positions.

4.3.3 Working of proposed PSO Algorithm

1. Create and initialize a $K - dimension$ swarm of P particles by choosing K CHs with residual energy higher than average energy of network for each particle.
2. Repeat
3. for each particle $i = 1,2,3 \dots \dots \dots.., P$ do
4. if $F(X_i) < lbest_i$ then
5. $lbest_i = X_i$
6. end

7. if $F(X_i) < gbest$ then

8. $gbest = X_i$

9. end

10. end

11. for each particle $i = 1,2,3 \ldots, P$ do

12. update velocity V_i using equation 4.1

13. update position vector X_i using equation 4.2

14. apply $\bigoplus_{NW}$ operator to updated position

15. end

16. until the maximum number of iteration reached

4.3.4 Working of proposed EEHCP Protocol

4.3.4.1 Optimal Number of Clusters

The optimal number of clusters K_{opt} can be found using simple analysis as in [105]. Let us assume an area of $M\ X\ M$ square meters with base station situated at the center and N sensor nodes are distributed over this area. The energy dissipated in the cluster head node assuming its distance from BS less than $d_{0,}$ during a round can be given by the following formula:

$$E_{CH} = \left(\frac{N}{k} - 1\right) L.E_{elec} + \frac{N}{k}.L.E_{DA} + L.E_{elec} + L.\varepsilon_{fs}.d_{toBS}^2 \qquad (4.9)$$

Where k is the number of clusters, E_{DA} is the data aggregation cost of a bit per report to the base station, and d_{toBS} is the average distance between the cluster head and the base station.

The energy used in a non-cluster head node is equal to:

$$E_{nonCH} = L.E_{elec} + L.\varepsilon_{fs}.d_{toCH}^2 \tag{4.10}$$

Here d_{toCH} is the average distance between a cluster member and its cluster head.

Thus energy dissipated in a cluster per round:

$$E_{cluster} \approx E_{CH} + \frac{N}{k}.E_{nonCH} \tag{4.11}$$

The total energy dissipated in the network is equal to:

$$E_{tot} = L(2.N.E_{elec} + N.E_{DA} + \varepsilon_{fs}(k.d_{toBS}^2 + N.d_{toCH}^2) \tag{4.12}$$

According to HEED [88]:

$$\mathrm{d}_{\mathrm{toCH}}^2 = \int_{x=0}^{x=x_{max}} \int_{y=0}^{y=y_{max}} (x^2 + y^2)\rho(x,y)dxdy = \frac{M^2}{2\pi k} \tag{4.13}$$

and

$$d^2{}_{toBS} = \oint_A \sqrt{x^2 + y^2}\,\frac{1}{A}.dA = 0.765\,\frac{M}{2} \tag{4.14}$$

Differentiating E_{tot} with respect to k and equating to zero, the optimal number of constructed clusters can be found:

$$K_{opt} = \sqrt{\frac{N}{2\pi}}\ \frac{M}{d_{toBS}^2}\ \sqrt{\frac{\epsilon_{fs}}{\epsilon_{mp}}} \tag{4.15}$$

4.3.4.2 Network Deployment Model

Most of the routing protocols for heterogeneous WSNs consider the random deployment of sensor nodes in the network space but to take advantage of heterogeneity present in the network in term of energy level we partitioned the network space into two zones: α- zone and β - zone.

This protocol assumes three levels of heterogeneity with three types of nodes: α-nodes, β-nodes and ω-nodes. α-nodes have the highest energy among all the nodes, β-nodes have energy less than α-nodes but greater than ω-nodes and ω-nodes are normal nodes. We assume that a small fraction of total nodes are α or β-nodes while most of the nodes are ω-node. As in SEP, we consider that the base station is not mobile and coordinates of the base station are known. The area in network space near the base station is β-zone and β-nodes are deployed in this region while the α-nodes are deployed in the farthest region, i.e. α-zone and ω-nodes are randomly deployed in the whole network space.

4.3.4.3 Cluster Head Selection

Let us assume that A is fraction of total nodes that are α-nodes and B is the fraction of total nodes that are β-nodes. The energy of an α-node is α times higher than a ω node while the energy of a β-node is β time higher than a ω node. If the initial energy of a ω-node is E_0 then the total energy of all the nodes will be:

$$E_{total} = N(1 - A - B)E_0 + N.A.E_0(1 + \alpha) + N.B.E_0(1 + \beta) \tag{4.16}$$

The optimal probability of cluster head selection in case of homogeneous network is given by [37]:

$$P_{opt} = \frac{k_{opt}}{N} \tag{4.17}$$

Where K_{opt} is optimal number of cluster heads as calculated above and N is total number of nodes in network.

Now optimal probability of a node to be cluster head on the basis of residual energy can be calculated as:

$$(P_\omega)_i = \frac{P_{opt}\, E_i(r)}{(1+A\alpha+B\beta)\bar{E}(r)} \tag{4.18}$$

$$(P_\beta)_i = \frac{(1+\beta)P_{opt}\, E_i(r)}{(1+A\alpha+B\beta)\bar{E}(r)} \tag{4.19}$$

$$(P_\alpha)_i = \frac{(1+\alpha)P_{opt}\, E_i(r)}{(1+A\alpha+B\beta)\bar{E}(r)} \tag{4.20}$$

Here $E_i(r)$ is residual energy of i^{th} node in r^{th} round and $\bar{E}(r)$ is the average energy in the r^{th} round.

Depending on the weighted probabilities the threshold values can be calculated as follows

$$T_\omega = \begin{cases} \dfrac{P_\omega}{1-P_\omega\left(1-r \bmod \frac{1}{P_\omega}\right)} & if\ \omega \in G'' \\ 0 & otherwise \end{cases} \tag{4.21}$$

$$T_\beta = \begin{cases} \dfrac{P_\beta}{1-P_\beta\left(1-r \bmod \frac{1}{P_\beta}\right)} & if\ \beta \epsilon G' \\ 0 & otherwise \end{cases} \tag{4.22}$$

$$T_\alpha = \begin{cases} \dfrac{P_\alpha}{1-P_\alpha\left(1-r \bmod \frac{1}{P_\alpha}\right)} & if\ \omega \epsilon G \\ 0 & otherwise \end{cases} \tag{4.23}$$

Where G, G' and G'' are the sets of α, β, ω-nodes that have not been the cluster head in last epoch respectively.

Each node generates a number in interval [0, 1] randomly. If this random number is less then corresponding threshold the node will become cluster head. Once the cluster head is selected, it broadcasts an advertisement message to all the nodes, a node that receives such message decides on the basis of received signal strength that to which cluster head it will associate for the current round.

4.3.4.4 Data Transmission Techniques

In this approach, we use two techniques for data transmission:

a) Single-hop direct transmission
b) Multi-hop transmission through cluster heads(CHs)

In single hop transmission, a node n near the base station will directly send the data to base station if

$$d_{n\ to\ BS} < {d_0}/{k'} \quad \text{and} \quad \text{Residual energy } E_n(r) \geq \bar{E}\,(r)$$

Here $d_{n\ to\ BS}$ is distance between n^{th} node and base station; k' is a parameter, which is used to control the single hop transmissions; $E_n(r)$ is residual energy of n^{th} node in r^{th} round; $\bar{E}\ (r)$ is average energy of whole network in r^{th} round.

If the above two conditions do not satisfy simultaneously, the node will send data to cluster head for further processing. Each cluster head creates a schedule based on TDMA and nodes send data to their cluster head at their respective time slots as described in the schedule. We used Dijkastra's shortest path algorithm to find the shortest route from a cluster head to base station through other cluster heads. For this, we have used the distance of nodes from each other and base station as weights.

4.4 Simulation Results and Analysis

For simulation, we assume a square network field of size 100m X 100m with 100 sensor nodes deployed uniformly in it. We assume that sink is at the centre of the field. Performance of proposed algorithm compared with LEACH and its popular variant LEACH-C. For the simulation of second proposed algorithm,the network field and number of sensors taken are same as in first but deployment is not uniform. 20m X 20m area centered at (50, 50) is normal-zone and the area of width 20m surrounding the normal zone is β-zone and β-sensors (b=no. of sensors) have been deployed in this zone randomly. Similarly, the remaining area of width 20m that surrounds the β-zone is α-zone and α-sensors (a=no. of sensors) are deployed in the α-zone remaining ω-sensors are deployed randomly in whole network field. MATLAB is used for the simulation, using different values of α, β, A and B. Simulation parameters are shown in Table 4.1.

Table 4.1: **Simulation Parameters**

Description	Parameter	Value
Initial energy of normal nodes	E_0	0.5 *J*
No. of α nodes as fraction of total nodes	A	0.1, 0.2
No. of β nodes as fraction of total nodes	B	0.1, 0.2
Energy coefficient for α nodes	α	2
Energy coefficient for β nodes	β	1, 1.5
Data aggregation energy	E_{DA}	5 *nJ/bit/signal*
Electronic circuitry energy	E_{elec}	50 *nJ/bit*
Free space co-efficient	ε_{fs}	10 p*J/bit/m2*
Multi-path co-efficient	ε_{mp}	0.013 *pJ/bit/m4*
Optimal percentage of CHs	P_{opt}	0.1
Initial Energy of α nodes	E_α	E_0 (1+α)
Initial Energy of β nodes	E_β	E_0 (1+ β)
Data packet size	P_{pkt}	30 bytes
Total no. of nodes	N	100

Fig. 4.1, 4.2 and 4.3 shows the number of alive nodes in each round of LEACH, LEACH-C and proposed protocol PSOBC for different positions of base station. Simulation results show a considerable improvement in lifetime of network.

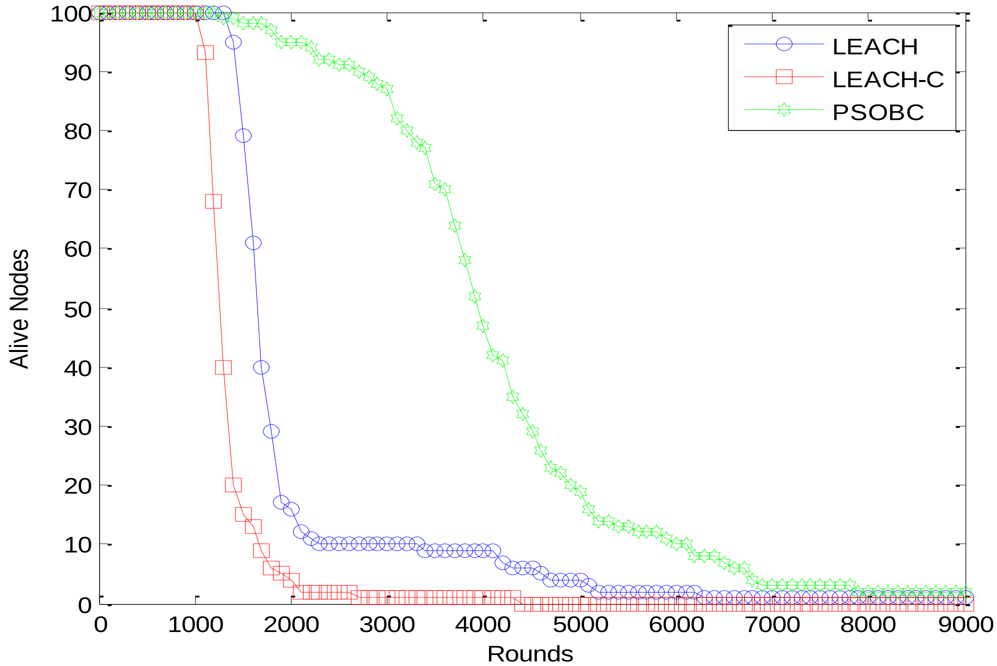

Fig. 4.1: Alive Nodes per Round for BS position (50, 0)

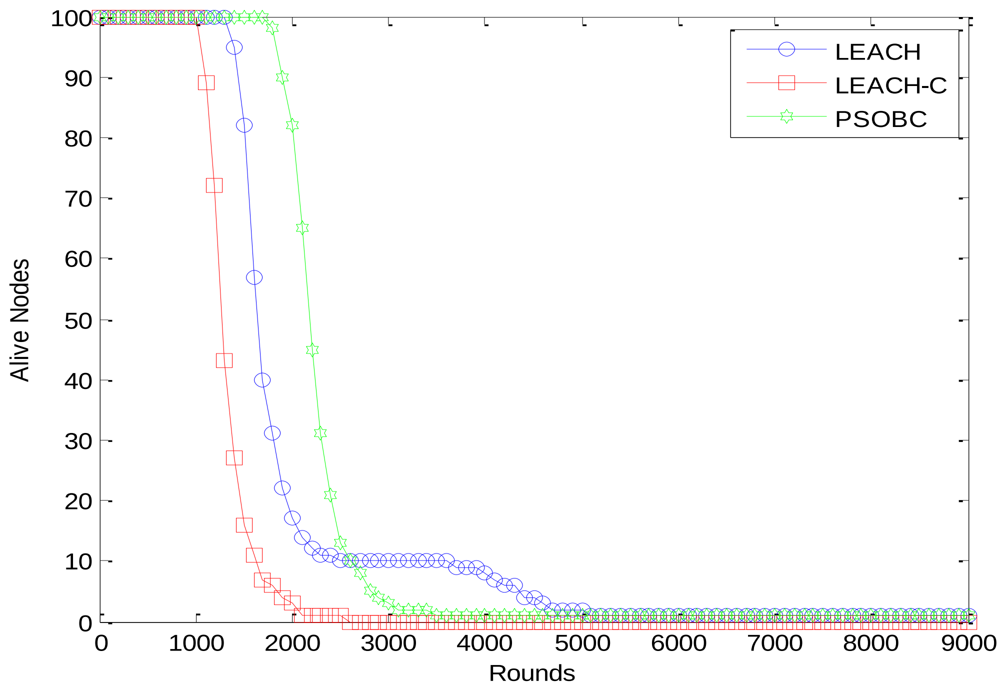

Fig. 4.2: Alive Nodes per Round for BS position (0, 50)

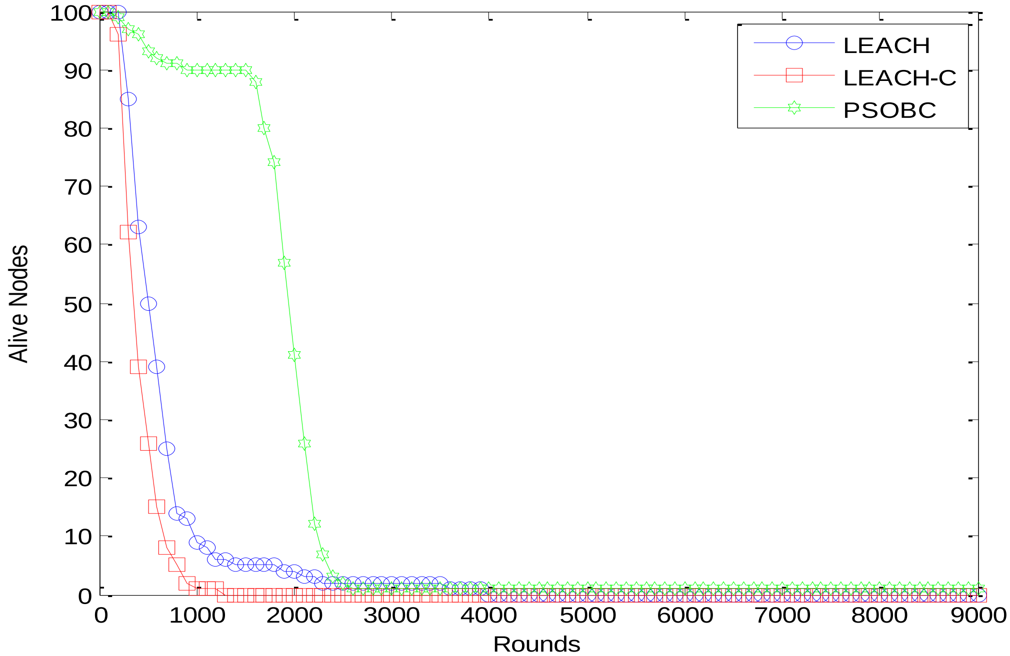

Fig. 4.3: Alive Nodes per Round for BS position (200, 200)

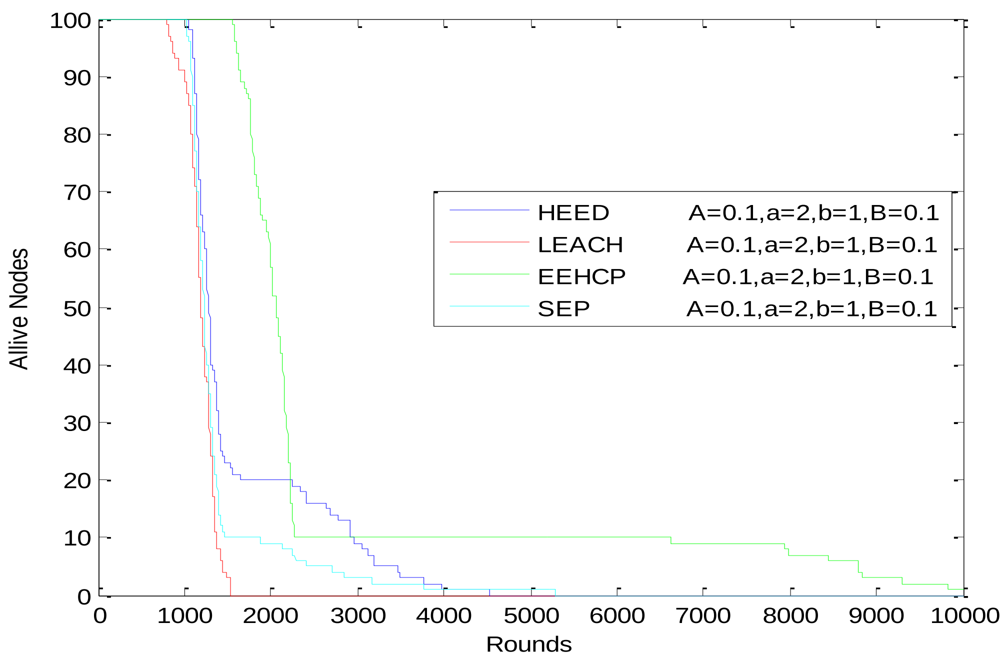

Fig. 4.4: No. of Alive Nodes vs Round

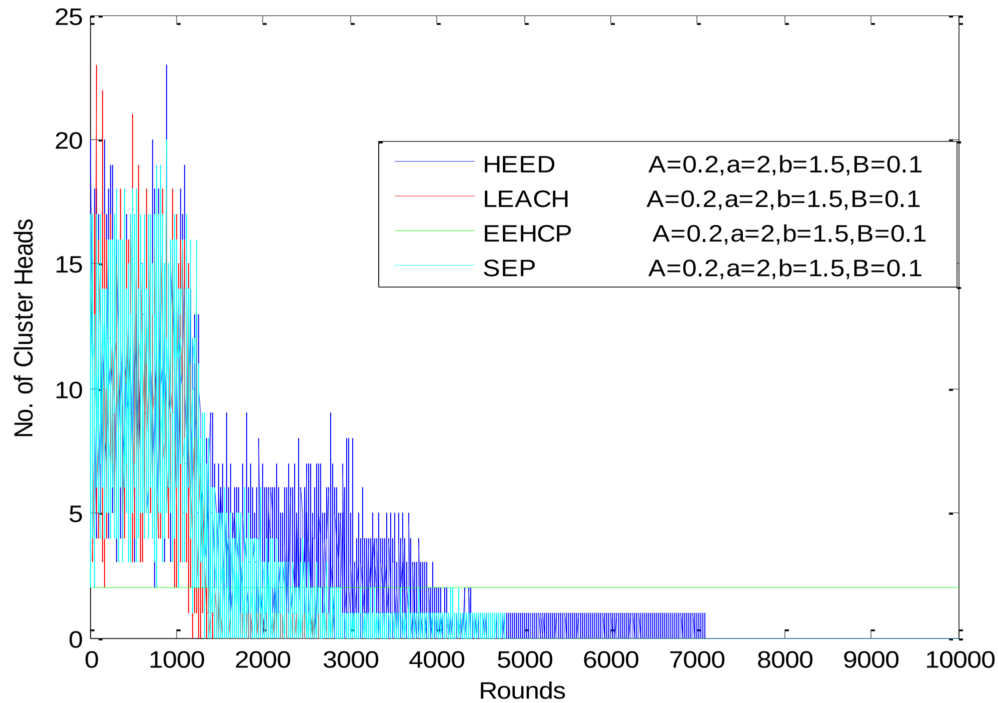

Fig. 4.5: No. of CHs vs Round

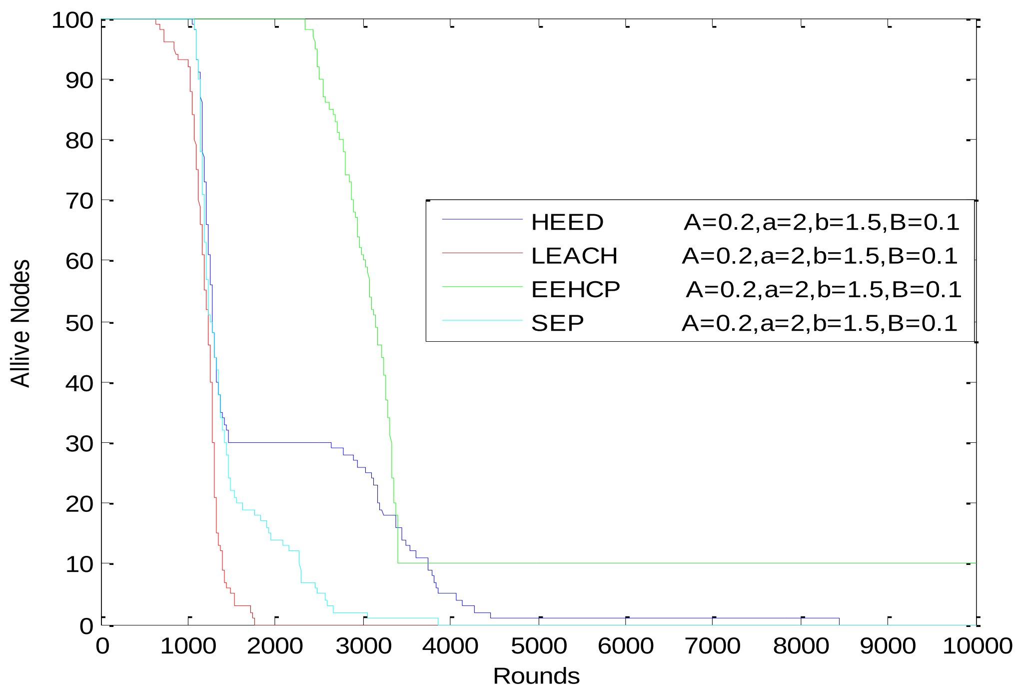

Fig. 4.6: No. of Alive Nodes vs Round

The above figures 4.4-4.6 shows a comparison of HEED, LEACH, SEP and EEHCP for different numbers and energy levels of α, β and ω nodes. In each case, EEHCP outperforms LEACH and SEP and HEED.

4.5 Conclusion

In this work, we proposed a PSO based solution to clustering problem in which a new operator is defined and used it with PSO algorithm to make it work with discrete search space. Simulation results show a considerable increment in Network lifetime as compared to LEACH and LEACH-C. In the second approach, hybrid cluster head selection mechanism applied which uses the heterogeneity in sensor node for an intelligent deployment of nodes in the network and the residual energy of nodes in particular data transmission round to weight the optimal probability of cluster head selection. The simulation result shows that proposed algorithm outperforms SEP, LEACH and HEED. It also prolongs the stability period and Network throughput.

Published work related to this chapter:

Rajesh K. Yadav, Daya Gupta and D.K. Lobiyal, "Energy Efficient Clustering Approach for Data Aggregation and Fusion in Wireless Sensor Networks," ***International Journal of Control Theory and Applications***, Vol.10, No. 13, pp. 241-253, 2017. [Scopus Indexed].

CHAPTER 5

DYNAMIC POSITIONING BASED AGGREGATION

As the utility of WSNs is being realized, in hindsight the various challenges related to sensor networks are coming to the fore, one such issue is the dynamic deployment of the sensor nodes in a monitoring area. This chapter presents a modified artificial bee colony algorithm that is used to deploy the mobile sensors with the aim of increasing the coverage area of the network in turn reducing redundancy and improving the performance of the network.

5.1 Introduction

With considerable technical advancement being achieved in the field of electromagnetic theory and communication theory, wireless sensor networks (WSNs) are being increasingly used in diverse applications in research, business, and surveillance domains. The WSNs can be made up of homogenous sensors or heterogeneous sensors. WSNs constantly monitor the mission area detecting events of interest and processing the gathered information, which is then relayed up to the base station for further analysis. While deploying WSNs special consideration is given to the problem of coverage and energy usage, optimization of these problems are crucial for the lifetime of the network.

The positioning of sensors of a WSN directly influences resource management. A crucial issue faced in sensor resource management is that of determining the precise positions of the sensors in the mission area that reduces communication cost, overheads and has a high area coverage, as a result tackling the problem of dynamic deployment of WSNs becomes increasingly important.

Researchers have extensively studied the problem and have proposed different approaches to tackling the problem. These problems fall into two broad categories namely static and dynamic deployment. In static deployment, the position of a sensor once fixed, it cannot be altered throughout the lifetime of the network. On the other hand dynamic deployment approaches are full-bodied in comparison to their static equivalent. In dynamic deployment, the sensors are mobile and hence have the ability to alter their positions if need arises like in the case of a node failure, in which case the topology of the network needs to be altered to compensate for the loss. The precise location where a sensor should be positioned can be determined either randomly or deterministically, but as it has been shown that the problem of dynamic node deployment is NP-Complete hence deterministic approaches are less preferred.

5.2 Related Work

The efficiency of a deployed wireless sensor network in accomplishing its aims is governed by how efficiently we can position the sensors of the network. The most intuitive approach for node deployment is random deployment that is easy to achieve but on another side, such deployments fail to satisfactorily achieve the objectives like maximizing area coverage, minimization of redundancy and communication cost, fault tolerance etc.

5.2.1 Dynamic Deployment of Sensor Network using Artificial Bee Colony Algorithm

In ABC [60] algorithm authors mimics the foraging behavior of honeybees. The food sources in the ABC algorithm denote a solution in the search space. The fitness of a food source is determined by its nectar value. In ABC algorithm a colony of bees is used as the swarm, the swarm of bees is made up of three types of bees namely the employee bee, the onlooker bee and the scout bee. The bee that initially wanders in search of food is called the employee bee, the bees that wait in the dancing area to choose a food source that was previously found by an employee bee are called the onlooker bees, and lastly the bee that looks for new food sources to replace the abandoned food source is called the scout bee.

In [120] authors were the first to use the ABC metaheuristic to direct the movements of sensors in a sensing field for maximizing coverage area. The authors used a sensor network that comprised of both static and mobile sensors. The authors used a food source of the ABC algorithm to represent a deployment sequence that has the positions of the sensors in the mission area and the coverage of the deployment that was calculated using equation 5.1 as its nectar value. The results showed that the ABC algorithm outperformed the PSO algorithm by generating a deployment sequence that covered more area.

1	2	3	4			2S-1	2S
X_1	Y_1	X_2	Y_2	…	…	X_S	Y_S

Fig. 5.1: A food source of the ABC algorithm here X_i, Y_i represents the Cartesian coordinates of the sensor S_i

Artificial Bee Colony algorithm [60] draws its inspiration from foraging behavior of honeybees. In a bee hive there are three types of bees, the first type is the explorer.The job of the explorer bee is to wander around the hive in search of food, once a food source is found the bee memorizes its path from the hive, it also factors in the distance to be travelled to reach this source ,the amount of nectar available, the effort required to reach the food source etc. once a food source is found the then returns to its hive, where the second type of bee is waiting. The role of the second type of bees is to exploit the food sources found out by explorer bees. The explorer bees communicate with these bees and inform them about what they have found, the quality of available food sources are graded by these bees and then decide on to which food source to exploit. The area in the hive where these activities take place called dancing area. Each explorer bees moves around in a random fashion, a movement which the authors termed as “waggle dance”. It is generally noticed that the better quality of the food source the longer the explorer bee dances. Lastly, there is the third type of bees whose sole responsibility is to wander around in hunt of new unexplored food sources, which are to be used as replacement for those food sources that have been exhausted or due to external factors no longer remain as profitable as before. The authors classified the bees into three categories namely Employee bees, Onlooker bees and Scout bees.

The employee bee carries out a local search around a found food sources in search of better food sources. The local search around a food source is modeled using equation 5.1.

$$X_{ij} = Y_{ij} + \varphi * \left(Y_{kj} - Y_{ij}\right) \tag{5.1}$$

Where X_{ij} is the new food source, Y_{ij} is the current food source, "k" is a real number that is not equal to "i" and "φ" is uniformly distributed random number in the range (-1,1). The random number "k" determines the magnitude of the movement from the current position where as "φ" determines the axis of the coordinate systems along which the movement takes place.

Once a new food source is discovered around an existing food source, the quality of both sources is evaluated and a greedy selection is made between the two. The fitness of all food sources discovered is normalized in the range between 0 to 1, using equation 5.2. The "onlooker bee" is used for exploitation of found food sources as well as searching new sources.

$$P_i = \frac{fitness_i}{\sum_{j=1}^{No.\ of\ food\ sources} fitness_j} \tag{5.2}$$

5.3 Coverage calculation strategy and fitness calculation model

5.3.1 Grid based coverage calculation strategy

In our work, we have employed the grid-based coverage calculation strategy wherein instead of covering the entire area we try to cover as many gird points as possible. The fitness of a deployment is measured in terms of the number of such grid points covered. In this approach, the entire area was divided into square grids of equal size and the four endpoints of each of these grids were considered as grid points. The coverage is measured as the total number of such grid points covered by all the sensors divided by the total number of grid points in the mission area i.e.

$$Coverage = \frac{\sum_{j=1}^{n} \mathrm{Cj}}{\text{Total number of gird points}} \; x \; 100\% \qquad (5.3)$$

Where C_j = number of grid points covered by sensor j.

5.3.2 Sensor Detection Model

There are two commonly used detection models for calculating the area coverage in wireless sensor networks namely the binary model and the probabilistic model. The binary model only cares about the detection range of the sensor and the Euclidian distance between the sensor and the point under consideration. It assumes that if the point to be sensed is within a sensor's sensing radius then the sensor with absolute certainty is able to sense the point. The binary model is expressed as following:

Consider there are m sensors in the deployment scheme and each having a radius r , if a sensor S_i is positioned at a point, $P(a, b)$ and the Euclidian distance between the sensor

S_i and a point X be denoted by $D(S_i, X)$ .The binary model is mathematically expressed using equation 5.4 .

$$C_{ab} = \begin{cases} 1, & D(S_i, X) < r \\ 0, & otherwise \end{cases} \tag{5.4}$$

The probabilistic model, on the other hand, states that the probability of detection of two points within the sensing radius of a sensor S is not always the same. The probabilistic model introduces a notion of uncertainty in the detection of a point if the Euclidian distance between the sensor and the point is greater than a distance denoted by r_{eff} . The model is expressed using the equation 5.5.

$$C_{ab} = \begin{cases} 0 & if\ r + r_{eff} \leq D(S_i, X) \\ e^{\frac{-\gamma_1 \alpha_1{}^{\beta_1}}{\alpha_2{}^{\beta_2}} + \gamma_2} & if\ \ r - r_{eff} < D(S_i, X) < r + r_{eff} \\ 1 & if\ D(S_i, X) \leq r - r_{eff} \end{cases} \tag{5.5}$$

Here γ_1, β_1, β_2 are measuring parameters, $\alpha_1 = r_{eff} - r + D\ (S_i, X)$

and

$\alpha_2 = r_{eff} + r - D\ (S_i, X)$ and γ_2 is the disturbing factor, r_{eff} is the detection uncertainty range.

Here,we have adopted the probabilistic model of detection. If an area X_{ovp} is covered by k_{ovp} sensors then a measure of the coverage is given by equation 5.6.

$$C_{ab}\ (X_{ovp}) = 1 - \prod_{S_i \in S_{ovp}} (1 - C_{ab}(S_i)) \tag{5.6}$$

The fitness of a particular coverage can be calculated with respect to a coverage threshold (C_{thres}) as in equation 5.7.

$$C_{ab}\ (X_{ovp})\ >=\ C_{thres} \tag{5.7}$$

5.4 Proposed work

A potential area of improvement in the ABC is in its local search. The ABC algorithm uses the equation 5.1 for exploring around an existing solution (local search) and equation 5.2 for calculating its normalized fitness value in the interval (0, 1). Both the employee bee as well as the onlooker bee use the same equations for finding new solutions. Researchers have studied the local search characteristic of the algorithm and have suggested multiple approaches for improving searching capability of the algorithm; their study exposed two areas where the local search falls short:

- The undirected or haphazard search for new food sources around old food source and the way it balances exploration and exploitation.

One approach suggested by researchers for overcoming the first shortcoming of the ABC is to modify the local search so that instead of randomly searching around an old solution, the search for new solutions be guided along the global best solution found till that instant. This approach derives its roots from another popular swarm intelligence meta heuristic called particle swarm optimization. In PSO while swarm of particles is exploring the solution space, each particle periodically recalibrates its movement as guided by two parameters namely $pBest$ and $gBest$. The particle factors in its own past best position and the position of the fittest particle in the swarm to adjust its velocity and position in the search space.

The local search of conventional ABC algorithm modified to take into consideration the current global best solution and try to search for new solution in its direction rather than randomly looking around the solution under consideration. The local search guided by global best is expressed by the following equation

$$X_{ij} = Y_{ij} + \varphi * (gBest_j - Y_{ij}) \tag{5.8}$$

The results of their experiments using the modified local search in the canonical ABC algorithm. The global best guided local search expressed in equation 5.9.

$$X_{ij} = gBest_j + \varphi * (Y_{kj} - Y_{ij}) \tag{5.9}$$

Where, $gBest_j$ represents the j^{th} parameter of the global best solution found so far, Y_{kj} and Y_{ij} represent the current solution under consideration and a randomly generated solution around the current solution respectively and φ is a random number ranging from -1 to 1.

- The second shortcoming of the ABC is in the way it balances exploration and exploitation phase.

The studies done in the past pointed out the fact that the local search of ABC favoured exploration to exploitation, which generally results in valuable domain knowledge gathered by the swarm of bees being thrown away. The solution to this problem is the use of variable length step size.

In this work, we incorporated a memetic search [162] for balancing the exploration and exploitation. Memetic search internally uses the golden section search to determine the step size. The golden section search cuts down the solution space in accordance to the fitness of the solutions by updating two parameters namely ‘a’ and ‘b’. it works as follows:

```
while (abs(a-b) < ϵ)
   F_S1 = b – ( b - a ) * Ψ;
   F_S2 = a + ( b – a ) * Ψ;
   if ( Obj(F_S1) < Obj(F_S2) )
         b = F_S2;
   else
         a = F_S1;
```

The parameter Ψ, is a constant value and is called the golden ratio. The value of Ψ is 0.618. Obj (F_{S1}) is the fitness of the solution generated by the objective function, Obj () with step size F_{S1}. Obj (F_{S2}) is the fitness of the solution generated by the objective function, Obj () with step size F_{S2}. ϵ is a constant, which defines the ending point of the algorithm, the value of ϵ, is 0.01.

The memetic search introduces a new parameter called the “perturbation rate’ which controls the modification of i^{th}dimension of the newly generated solution. To summarize, the proposed hybrid local search contains the following computations.

Pseudo code for hybrid local search
For a given food source Y_{ij} , generate new solutions namely, U_{ij} using the local search of conventional ABC algorithm using equation 5.1 V_{ij} using the guided best search using equation 5.9 W_{ij} using memetic search Make a greedy selection between U_{ij}, V_{ij}, W_{ij} based on their fitness value.

$$X_{ij} = \max(fitness(Uij), fitness(Vij), fitness(Wij))$$

Another possible scope for improvement in the functioning of ABC algorithm is to use the fittest food sources food to generate new food sources. The ABC algorithm relies on a random search that is based on selecting a random food source and a random dimension for new food sources and hence does not make use of the fittest solutions that have already been discovered. A crossover operator similar to the one used in Genetic Algorithm (GA) has been incorporated before send scout bee phase in ABC. The genotypes for the crossover operator is the food source as shown in Fig. 5.1. For the crossover operator, a population of size equal half of number of food sources is generated using tournament selection.

For generating the population, we sorted the food sources (let say there were 'F' food sources in total) in decreasing order, in accordance to their fitness values and then we selected 3 random food sources out of them and included the fittest food source amongst the three, in the population. Once the population was created, a fixed amount of food sources selected from the median of the sorted food sources for replacement using crossover. The number of food sources selected for replacement is the product of crossover probability and the number of food sources.

For each of the selected food source pick two parents randomly from the population and after generating two random crossover points we crossed the two parents to produce two new solutions (food sources) as shown in Fig. 5.2. The selected food source is then replaced by the best among the two newly created food sources and itself.

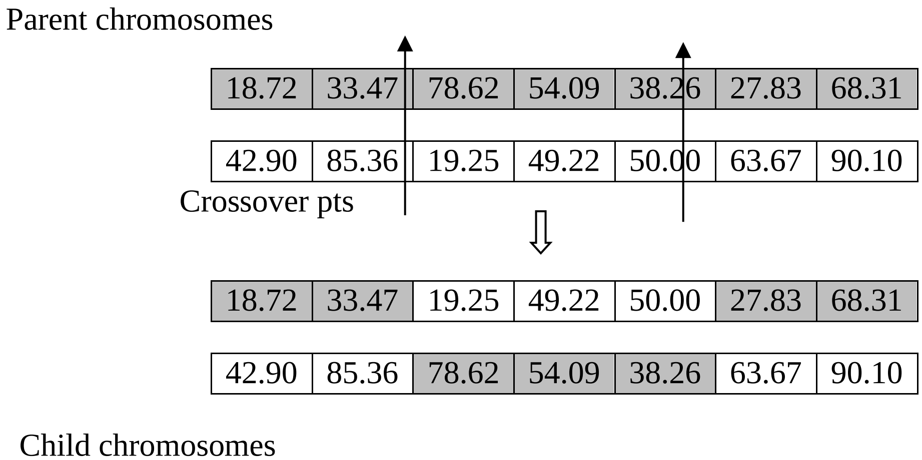

Fig. 5.2: **Two-point crossover**

Pseudo code for dynamic node deployment using proposed algorithm

1. Initialize the Control Parameters of the ABC Algorithm: initialize the number of food sources, number of bees = colony size/2, limit etc., randomly generate the positions of the sensor nodes, perturbation rate = 0.4, ϵ = 0.01, set the crossover_probability = 20%.
2. Set Iteration = 0

   Initialize the food source positions.

   Evaluate the nectar amount (fitness) of food sources.
3. Repeat Until (cycle is not equal to MAXCYCLE)

   for each employee bee do

   a) Search the neighbourhood of the food source for new solutions using hybrid local search

   a) Check if X_{ij} is within the bounds of the monitoring area.

   b) Evaluate the fitness of the new food source using equation 5.6.

   c) Make a greedy selection between old solution and the new solution.
4. Compute the probability $Prob_i$ of the solution using equation

$$Prob_i = \frac{0.9 * fitness_i}{fitness_{best}} + 0.1$$

5. For each onlooker bee do

a) Generate a random number r ϵ (0, 1) and select a food source depending on the value of $Prob_i$ and r.

b) Look for a new food source in the neighbourhood using hybrid local search

c) Check if the new solution is within the bounds of the area

d) Evaluate the fitness of the new food source using equation 5.6.

e) Make a greedy selection between the old solution and the new solution.

6. Select a fixed number of food sources (N) for replacement using the formula

$$N = crossover\_probability * number\ of\ food\ sources$$

a) Generate the mating pool for crossover operation using tournament selection.

b) For each selected food source do

i) Make a random selection of two food sources that will serve as parent genotype.

ii) Generate two random crossover points and swap genes as in Fig. 5.2

iii) Make a three-way greedy selection amongst the selected food source and the two newly created food sources as a candidate for replacing the selected food source.

7. Iterate through the trial array of food sources

if (the trials for a food source is greater than max trials)

a) Replace the food source with a randomly generated food source.

b) Set trial = 0.

8. Memorize the best solution found so far.

9. cycle = cycle +1.

5.5 Experimental Setup and Result Analysis

The experimental setup is as follows, we modeled a sensor network with 45 mobile sensors. Each sensor has a sensing radius of 7m, the detection error r_{eff} is 3.5m, the monitoring region is a square area of $10000m^2$, the value of the parameters of the probabilistic model are as follows: γ_1=1, γ_2=0, β_1=1, β_2=0.5,C_{thres} = 0.9. The crossover probability is 0.2 and the tournament population size is 20. The colony size is 80 and the number of food sources is 40. The scenario of a random deployment of sensors is run 29 times with each run having 1000 cycles.

For comparing the performance of the proposed algorithm with the ABC, in each iteration we generated random food sources that served as input for both the algorithms. Both algorithms worked on the same input set and produced results in the form of a 1x2S matrix, where ‘S’ is the number of sensors as shown in Fig. 5.1.

The Fig. 5.3 shows the initial random deployment of sensors. The grid points are denoted in blue colour. The area was divided into squares of 5x5 dimensions, resulting in a total of 441 grid points, in the figure however, 121 gird points (end points of squares of 10x10 dimensions) are shown rather than 441 for the sake of improved visibility. The points marked in red are the center points of the sensors whose sensing radius is pictorially represented by a circle.

5.5.1 Performance Metric

The “Maxcycles” is the parameter, which is used as metric for evaluation of the proposed algorithm and ABC algorithm. This parameter denotes the stipulated number of trials that the algorithm carries out for finding better solutions starting from a randomly generated solution that was fed to it as an input.

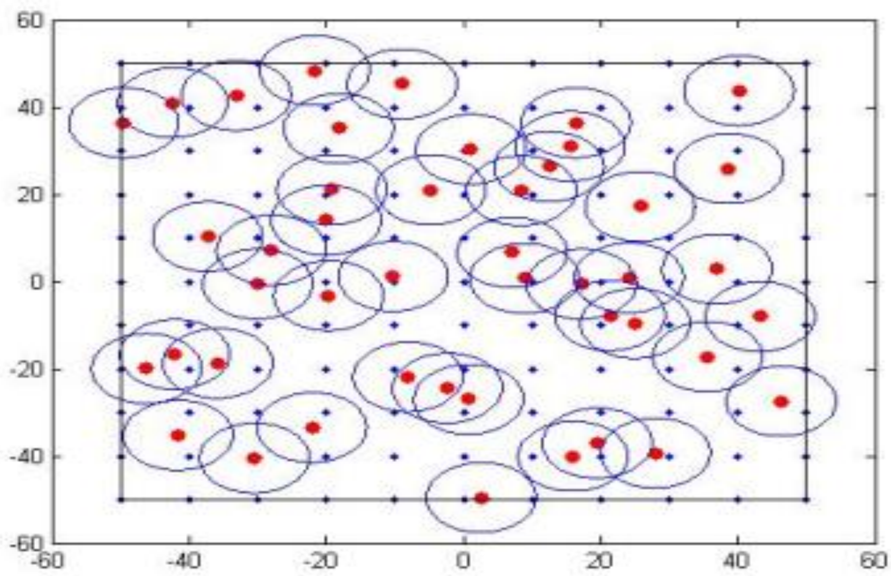

Fig. 5.3: **Initial deployment sequence**

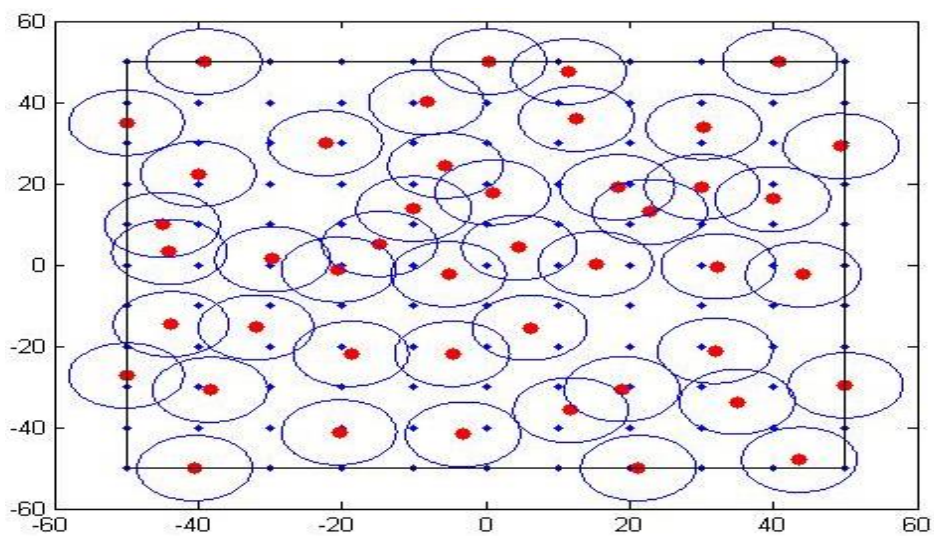

Fig. 5.4: **Best deployment sequence with conventional ABC**

Fig. 5.4 shows the best deployment returned by the ABC algorithm starting with the deployment in fig.5.3 as input, the coverage of this deployment is 84.58%. The fig. 5.5 shows the best deployment of the proposed algorithm, the coverage of this deployment is

89.79 %. The fig. 5.6 – 5.9 shows the best deployment sequence found by ABC at various cycles. The deployment at 50th, 100th, 500th and 1000th cycle is shown in figures 5.6, 5.7, 5.8 and 5.9 respectively.

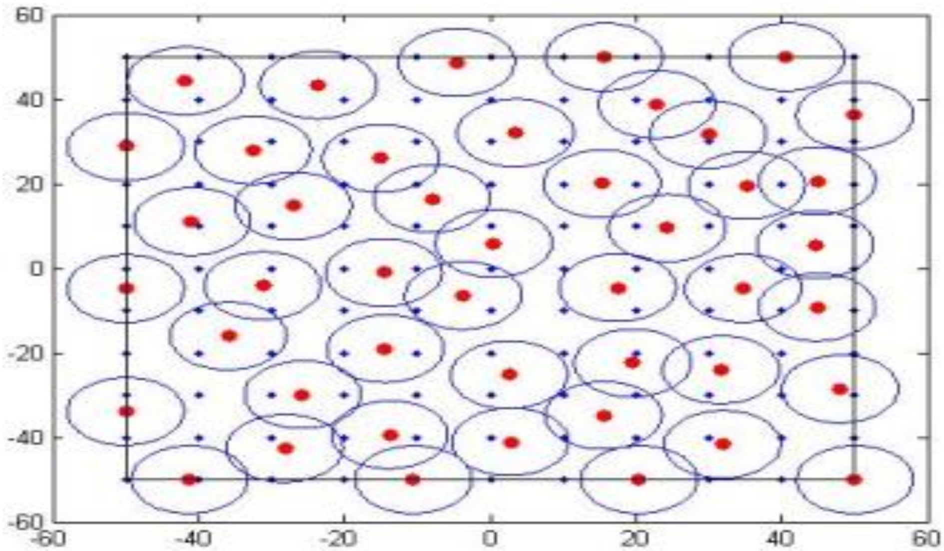

Fig. 5.5: Best deployment sequence of proposed algorithm

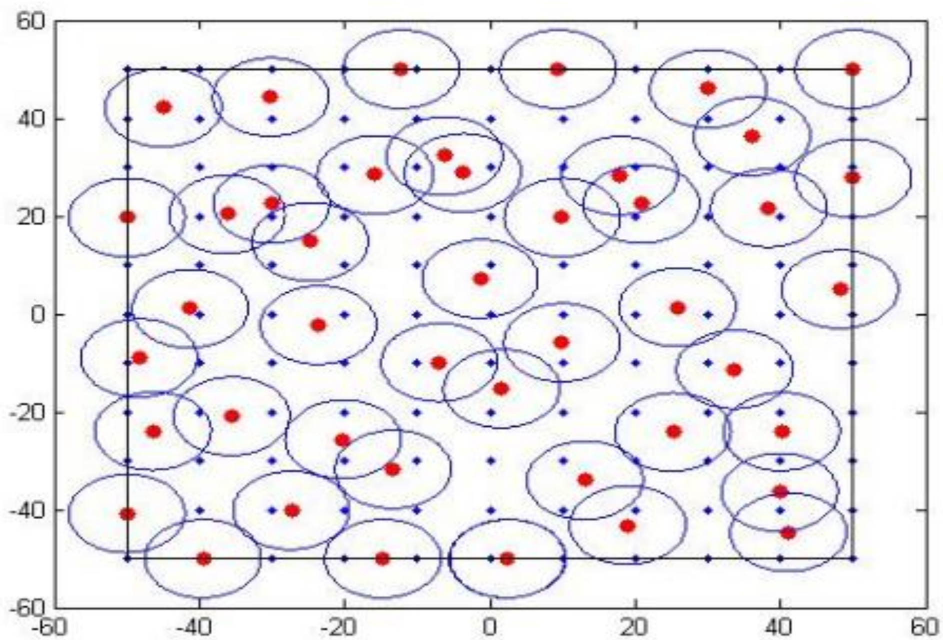

Fig. 5.6: Best deployment sequence by ABC at 50th cycle

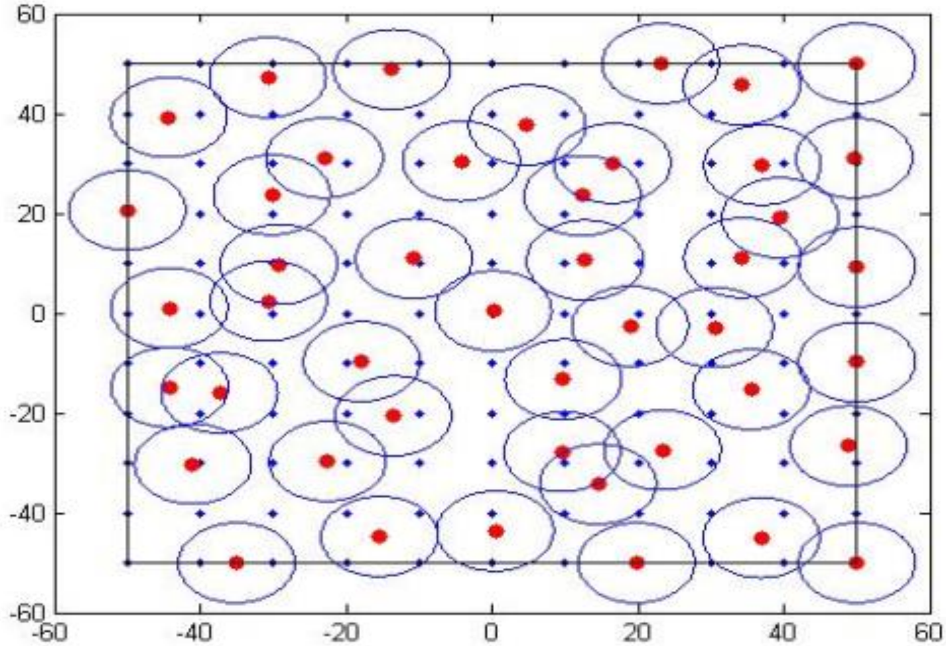

Fig. 5.7: Best deployment sequence by ABC at 100^{th} cycle

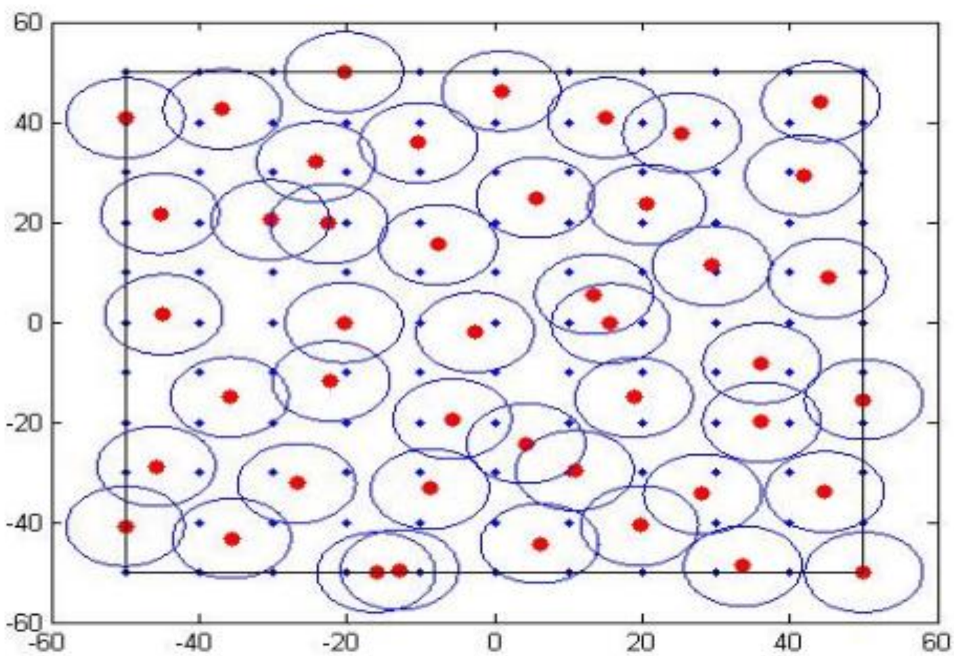

Fig. 5.8: Best deployment sequence by ABC at 500^{th} cycle

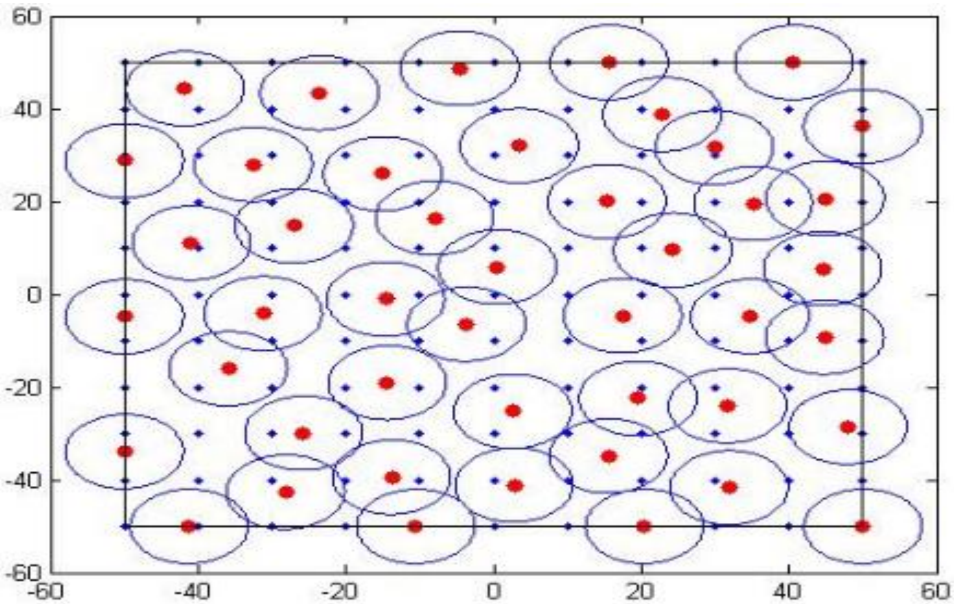

Fig. 5.9: Best deployment sequence by ABC at 1000th cycle

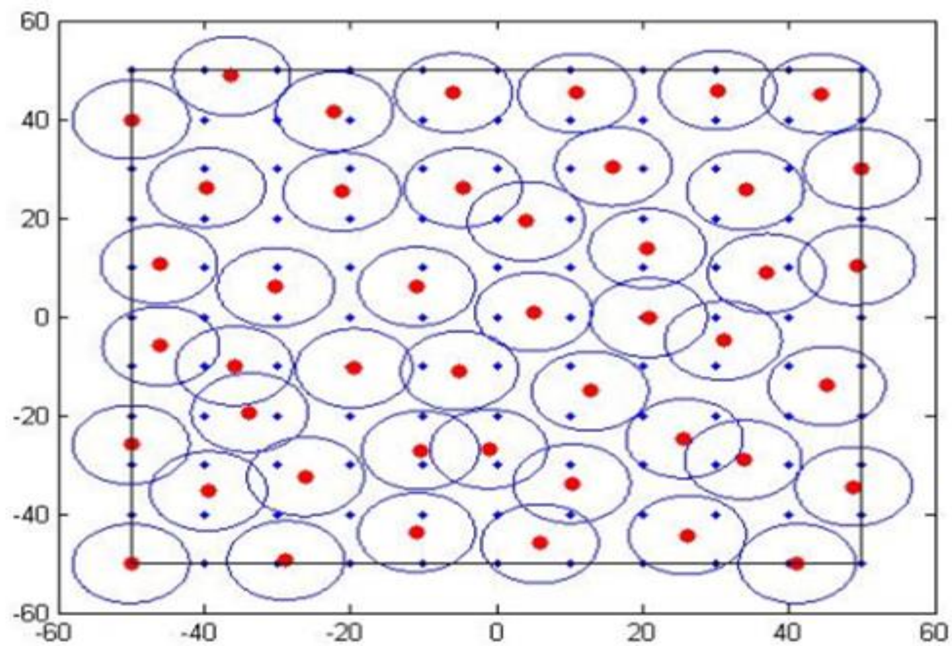

Fig. 5.10: Best deployment sequence by the proposed algorithm at 50th cycle

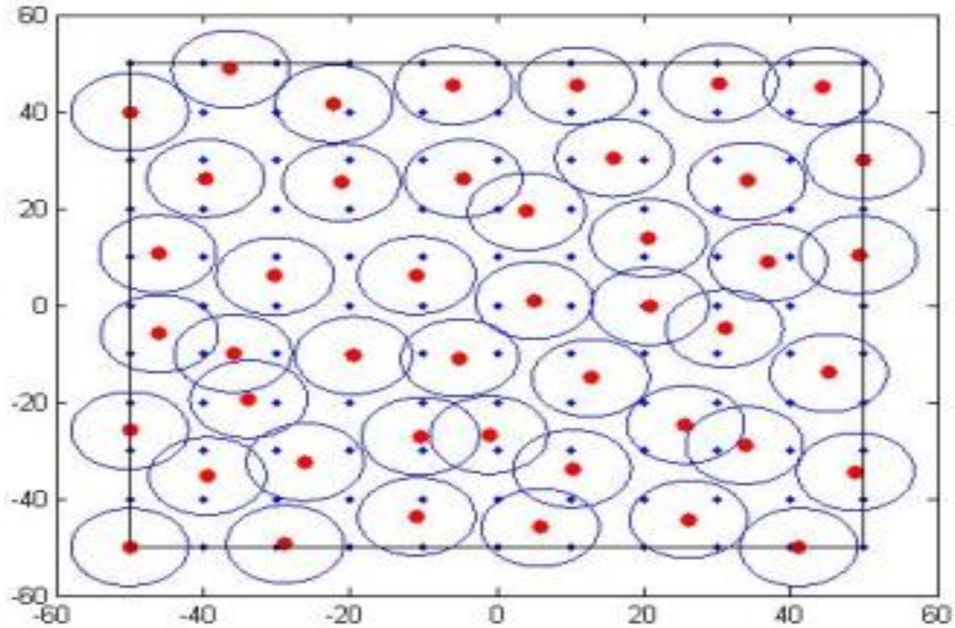

Fig. 5.11: Best deployment sequence by the proposed algorithm at 100th cycle

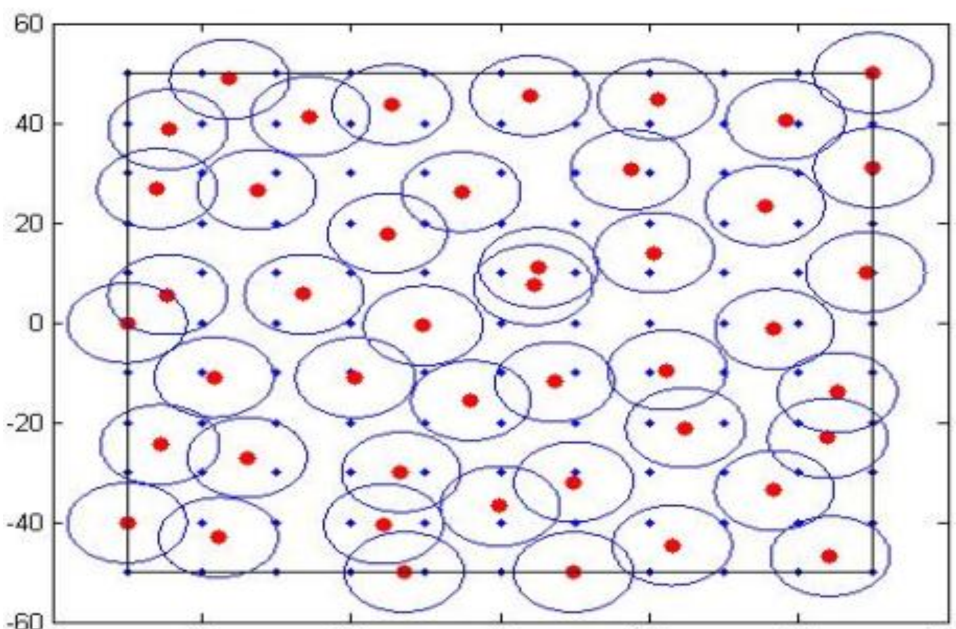

Fig. 5.12: Best deployment sequence by the proposed algorithm at 500th cycle

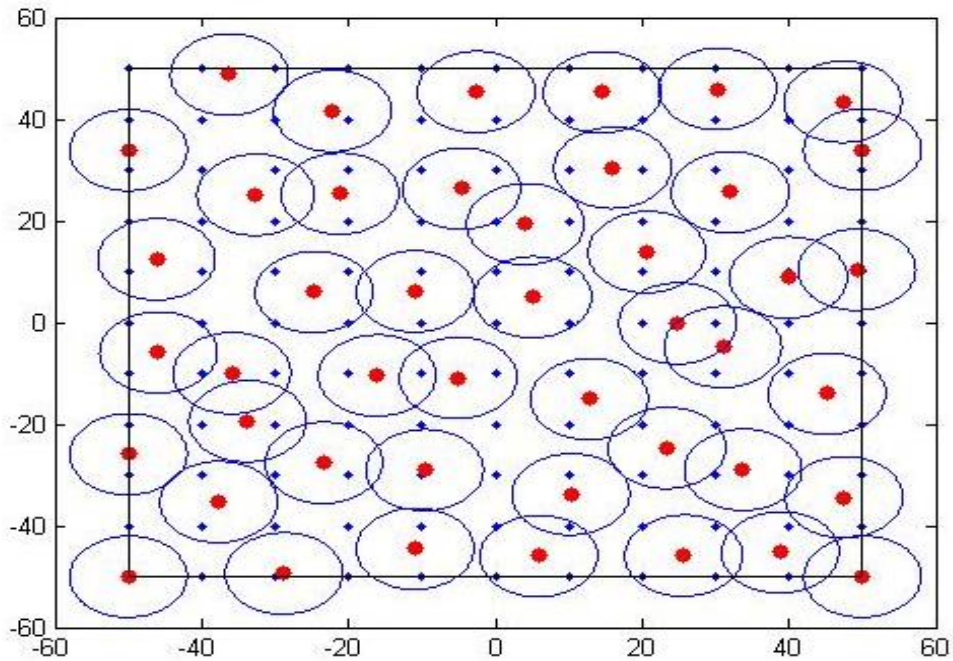

Fig. 5.13: **Best deployment sequence by the proposed algorithm at 1000th cycle**

The figures 5.10 - 5.13 show the best deployment sequence found by our proposed algorithm at various cycles. The deployment at 50th, 100th, 500th and 1000th cycle is shown in figures 5.10, 5.11, 5.12 and 5.13 respectively.

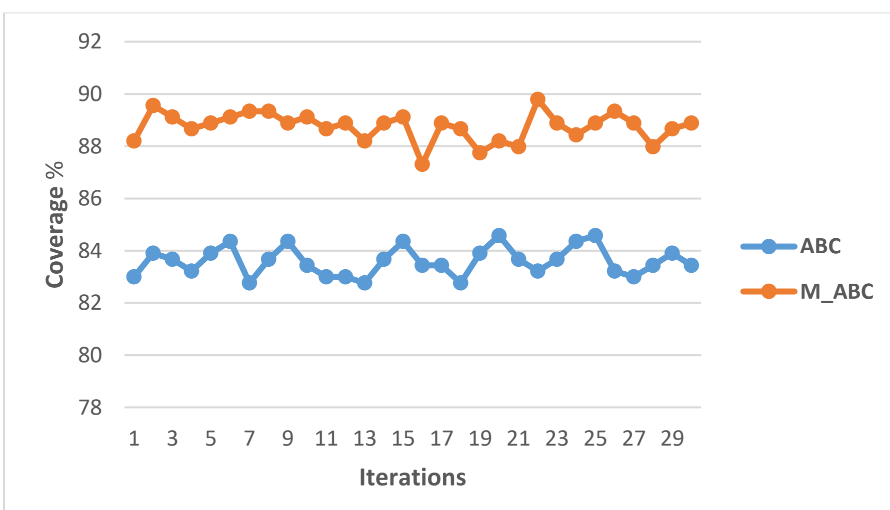

Fig. 5.14: **Comparison of ABC and Proposed (Modified-ABC)**

Fig. 5.14 shows the execution of artificial bee colony algorithm (ABC) and our proposed algorithm i.e. Modified ABC (M_ABC). The mean coverage of ABC for this experimental setup was 83.58% whereas as that of our proposed algorithm was 88.84%.

5.6 Conclusion

In this chapter, we have demonstrated an ABC based technique for dynamic positioning of sensor network. We addressed two deficiencies of the conventional ABC algorithm using a hybrid local search and a crossover operator. The results showed that the proposed algorithm outperformed the ABC algorithm by finding positioning sequences having higher coverage area hence minimizing redundancy.

Published work related to this chapter:

Rajesh K. Yadav, Daya Gupta and D.K. Lobiyal, "Dynamic Positioning of mobile sensors using Modified Artificial Bee Colony Algorithm in a Wireless Sensor Networks," ***International Journal of Control Theory and Applications***, Vol. 10, No. 18, pp. 167-176,2017. [Scopus Indexed].

CHAPTER 6

ROUTING BASED IN-NETWORK DATA AGGREGATION

Data aggregation is important in energy constraint mobile wireless sensor networks (MWSNs) which can make use of the energy of the sensor nodes efficiently and reduce the traffic in the network. Components of in-network aggregation include: routing protocols, aggregating functions and ways of representing the data. This chapter focused on a routing approach which exploits the principles of swarm intelligence inherent in ant colony optimization (ACO) for routing and the goodness of a path is evaluated by a set of fuzzy rules.

6.1 Introduction

Many applications of WSNs have stationary nodes and many others have mobile sensor nodes include live stock breeding and migration pattern tracking where sensor nodes are attached to animals, sensors are attached to patients to monitor their vital parameters and sensors are attached to unmanned aerial vehicles for surveillance or environmental monitoring.In all these applications the movement of the node is arbitrary just like in MANETs (Mobile Adhoc Networks). As there is no fixed infrastructure between wireless sensor networks for communication, routing becomes an issue in large number of sensor nodes deployed along with other challenges of these networks. In Mobile Wireless sensor network applications, AODV is the most suitable routing protocol due to its advantages of quickly adapting to dynamic link conditions, its requirement of low memory and processing overhead [143, 150].

Ant colonies, and more generally social insect societies, are distributed systems that, in spite of the simplicity of their individuals, present a highly structured social organization. Because of this organization, ant colonies can accomplish complex tasks that in some cases far exceed the individual capabilities of a single ant. The ant colony optimization is one of the most successfully proven swarm intelligence. It has been successfully applied in many difficult discrete optimization problems such as the traveling salesman problem [61, 153]. Fuzzy logic consists of a decision system approach, which works similarly to the human control logic. It provides a simple method to reach a conclusion from imprecise, vague, or ambiguous input information [151].

6.2 Related Work

Recently many researchers have been attempting to address this problem. Being attentive to structural constraints that have mobile wireless sensor networks such as sensor limitation energy, developing efficient algorithms for routing seems necessary [137, 152, 144]. In [140] authors proposed a novel routing approach using an ant colony optimization algorithm which uses artificial ants. Each ant chooses the next hop, moreover, the pheromone concentration amount attends to the node's remaining energy by this method, the ant selects a node with longer lifetime.

6.3 Proposed Work

The proposed routing approach is hybrid in nature and employs artificial ant routing packets for updating routing tables, which store pheromone values indicating the "goodness" of paths. When a source node needs to send data packets to a destination node (Sink node), but it has no routing information for the same, reactive ants are broadcasted from it towards the sink node. Broadcasting implies that the ants perform a breadth-first search from the source node towards the destination. All nodes except the destination (Sink node), drop duplicate ants based on sequence numbers. This allows multiple paths to be set up between the source and the destination, each being the shortest path between them for the given neighbor of destination present in the path, by virtue of breadth-first search.

Once the ants reach the destination, they are sent back towards the source retracing their paths to update the routing table entries by "depositing" pheromone values on the links between nodes. In general, when a path is used for routing packets, it will go down when any of the nodes in it run out of energy. Hence paths having comparatively low energy values of their nodes should be avoided. In order to evaluate the lifetime of a path, we consider the node in it having the minimum remaining energy among all the nodes in the path, for it is the most likely to go down first and break the path. In order to be energy efficient and hence prolong the lifetime of the network, lower pheromone values deposited on paths having lower values of energy of their minimum energy nodes, other factors remain same. Once the source node has established a path towards the destination, it proactively sends out ant packets, which follow the stored pheromone values to search for new paths and maintain existing ones.

6.3.1 Data Structures at each node

A number of data structures are maintained at each node. These are required for neighborhood maintenance, generation of proactive ants, and routing of packets. The routing information of a node i is represented in a pheromone table T^i.

Table 6.1: **Structure of Pheromone Table**

Neighbour	Destination	Pheromone value
n	d	ph_value

Each entry, $T^i{}_{nd} \epsilon R$ represents the pheromone value for a (Neighbour = n, Destination = d) pair, which represents the pheromone value on the link from node i to its neighbour n for routing packets to the destination d. Higher pheromone values indicate better path quality in terms of number of hops and path stability .

Neighbour table of a node i, N^i maintains all its current neighbours in the network. This is done via periodic broadcasts of Hello packets. This allows a node to maintain a 1-hop path to its neighbours in the pheromone Table 6.1.

Table 6.2: **Structure of Neighbour Table**

Neighbour	Expiry Time
n	t

Also associated with each neighbour node n, is an expiry time value-indicating till when the link between nodes i and n is assumed to be up. On receiving Hello Packet from node n, the expiry time for the link between nodes i and n is updated to a future value. Node i also periodically checks its neighbour table and removes entries whose expiry time is less than the current time.

The active destination table at a node i, D^i maintains a list of all the destinations for a node i. This table is used to randomly select a destination to which Proactive Ants are sent for discovering new paths and/or maintaining existing ones. The sequence number cache at a node i, $C^{\,i}$ acts as a look-up to avoid repeated broadcasting of the same generation of ants. Its structure is as follows:

Table 6.3: **Structure of Sequence Number Cache**

Source	Destination	Sequence Number
s	d	seq_num

Each entry, $C^i{}_{sd}$ represents the sequence number for a (Source = s, Destination = d) pair, which represents the highest sequence number of a Forward Ant packet received by node i which are sent over the network by source node s for destination node d. On receiving a forward ant packet sent by source node s for destination node d, node i looks up $C^i{}_{sd}$ and will only process the packet if it has a higher sequence number than the stored value and otherwise discards it.

6.3.2 Neighbourhood maintenance using hello packets

During the course of a communication session, a source node performs proactive path maintenance and exploration. It takes proactive actions to update the information about the currently used paths to the destination and tries to find new and better ones when possible. An important role in this process is played by hello packets. Hello Packets are short messages broadcast every T_{hello} seconds by all nodes. In proposed protocol T_{hello} set to 1 second.

Hello, messages are used in many existing protocols to allow nodes to detect which are their immediate neighbours. When a node i receives a hello message from a new node n, it can assume that n is its neighbour. After that, i expect to receive a hello message from n every T_{hello} seconds. After missing a certain number of expected hello packets i assume that n has moved away, and no longer considers it a neighbour. Here a node i allowed to miss a maximum of 2 hello packets from any of its current neighbours before erasing it from its neighbour table.

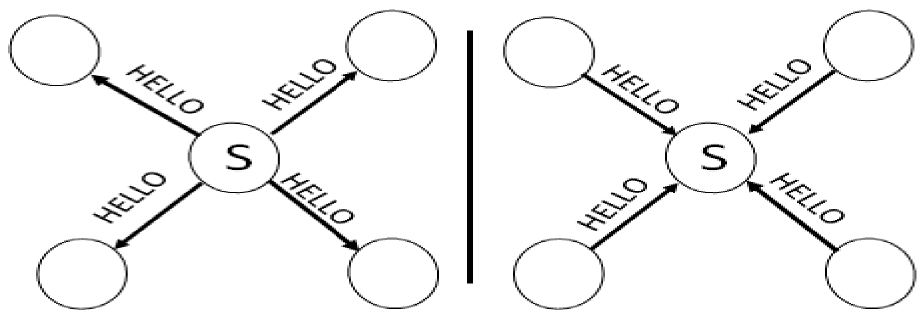

Fig. 6.1: **Hello packets for neighbourhood maintenance**

Here a node S is sending hello packets to its neighbour nodes. Therefore, S is listed in the neighbour table of the other nodes. Similarly, the four nodes in turn send hello packets in S, so S lists them in its neighbour table.

6.3.3 Ant Routing Packets

To update pheromone tables in proposed approach employs routing packets whose behaviour is modeled on the lines of ants exploring paths to a food source from their nest. Ants can be either forward or backward ants depending on whether they are traveling from a source to a destination node, or from a destination to a source node. They can be either reactive or proactive depending on whether they were sent out reactively or proactively. Hence there are 4 types of routing packets, namely Reactive Forward Ants (REFA), Reactive Backward Ants (REBA), Proactive Forward Ants (PRFA) and Proactive backward Ants (PRBA).

6.3.4 Reactive Forward Ant (REFA)

When a source node s starts a communication session with a destination node d, and it does not have routing information for d available, it broadcasts a REFA packet. The task of these ants is to discover a path to the destination node. At each node, the REFA is either unicasted or broadcasted, according to whether or not the current node has routing information for d. For getting the routing information the pheromone table is consulted. If routing information is available in the pheromone table, the ant chooses its next hop n for destination d with the probability P_{nd}:

$$P_{nd} = \frac{T_{nd}^{i}}{\sum_{j \epsilon N^{i}} T_{nd}^{i}} \tag{6.1}$$

To avoid repeated broadcasting of the same generation of ants, each forward ant carries a sequence number, and it is processed by a node only if the receiving node has not a seen higher sequence number forward ant for the source and destination pair. To assign sequence numbers to a packet, each node maintains a counter initially set to zero. Whenever a sequence number needs to be assigned to a routing packet, the sequence number of the packet is set to the current value of the counter and the counter is then incremented by one. The REFA also keeps a list of nodes it has visited in its “memory” to be able to backtrack to the source node in case it reaches the destination node. Upon arrival at the destination d, it is converted into a backward ant, which travels back to the source retracing the path stored in its memory. A REFA packet carries a number of fields that store information for routing the packet. The fields present in a REFA packet are summarized in the following Table 6.4:

Table 6.4: **Fields in a REFA Packet**

S. No.	Field Name	Purpose
1.	Ant Direction	Specifies whether the ant is a forward or a backward ant. It is set to "Forward" for a REFA packet
2.	Ant Type	Specifies whether the ant is a reactive or proactive ant packet. It is set to "Reactive" for a REFA
3.	Sequence Number	Specifies sequence number of the REFA packet
4.	Source	Specifies the address of the node which originated the REFA packet
5.	Destination	Specifies the address of the node to which the source node wants to discover a path
6.	Memory Size	Specifies number of entries in the memory of the REFA packet
7.	Memory	It is a list which stores address of nodes visited by the REFA packet

6.3.5 Reactive Backward Ant (REBA)

Upon arrival at the destination d, it the REFA packet converted into a REBA packet, which travels back to the source retracing the path stored in its memory. Its aim is to deposit pheromone on the links it has travelled which reflect the "goodness" of the path it is travelled to reach the destination such that better quality paths have higher pheromone values that poorer ones. To do so, after having arrived at an intermediate node i, the REBA calculates the number of hops between node i and the destination, and the minimum energy of the path (energy of node having minimum value of energy on that path) between node i and the destination. Both these values are used to update the pheromone table entry for node i. The update rule is a fuzzy function, which takes the number of hops and the value of minimum energy as an input and as an output that represents the pheromone update value for that link.

The value of ${T^i}_{nd}$ i.e. the pheromone value on the link from node i to its neighbour n for routing packets to the destination d is updated as follows:

$$T_{nd}^{i} = \alpha * T_{nd}^{i} + (1-\alpha) * \tau_{nd}^{i}, \qquad \alpha \epsilon [0, 1] \tag{6.2}$$

$$\text{Where, } \tau_{nd}^{i} = Fuzzy(N_i^d, E_i^d) \tag{6.3}$$

N_i^d = number of hops travelled from i to d and is defined recursively as follows assuming that the REBA arrives at node i+1 from node i

$$N_{i+1}^d = N_i^d + 1 \tag{6.4}$$

$$N_d^d = 0 \tag{6.5}$$

E_i^d = minimum energy of intermediate node in the path from node i to d and is defined recursively as follows assuming that the REBA arrives at node i+1 from node i

$$E_{i+1}^d = min(E_i^d, e_{i+1}) \tag{6.6}$$

$$E_{i+1}^d = e_d \tag{6.7}$$

Here e_i represents the energy of a node i.

After backtracking along the sequence of nodes listed in its memory, the REBA packet will eventually reach the source node and update its pheromone table entries for the destination node. The source node now has routing information available for the destination node.

Since multiple paths may be discovered by the REFA packets, the source node stochastically selects one among multiple next hops for the data packet. The source node now generates a next hop from its pheromone table in a manner similar to the case when a REFA packet is unicasted, which is given by equation 6.1. Therefore, in effect there are multiple paths set up between the source and the destination node, which makes the proposed protocol a multipath routing protocol. A REBA packet carries a number of fields, which store information for routing the packet. The fields present in a REBA packet are summarized in Table 6.5:

Table 6.5: **Fields in a REBA Packet**

S. No.	Field Name	Purpose
1.	Ant Direction	Specifies whether the ant is a forward or a backward ant. It is set to “Backward” for a REBA packet
2.	Ant Type	Specifies whether the ant is a reactive or proactive ant packet. It is set to “Reactive” for a REBA
3.	Source	Specifies the address of the node which originated the REFA packet corresponding to the REBA packet
4.	Destination	Specifies the address of the node to which the source node wants to discover a path
5.	Memory Size	Specifies number of entries in the memory of the REBA packet
6.	Memory	It is a list which stores address of nodes which need to be visited by the REBA packet to backtrack to the source node
7.	Hop Count	Specifies the number of hops visited by the REBA packet.
8.	Minimum Energy	Specifies the energy of the node having the minimum energy among the nodes visited by the REBA packet

6.3.6 Use of Fuzzy Logic for Pheromone Update

In the proposed protocol, fuzzy logic is used for updating the pheromone value of the links. This pheromone value is determined by a combination of two input fuzzy variables:

1. Number of hops: This is the number of intermediate nodes between the current node (node i) and the destination node (node d). Lesser number of hops gives a better path.
2. Minimum Energy: This is the remaining energy of the intermediate node having the minimum value of energy on the path between node i and destination d. A higher energy value corresponds to a more stable link and vice-versa.

The fuzzy variables, namely Pheromone Update, Number of hops and Minimum energy are defined as having linguistic values, low, moderate and high corresponding to low, moderate and high values, respectively of the corresponding variables and are defined by the following set of equations

Equation for the LOW region:

$$Y = \begin{cases} 1 & , 0 < x < a \\ \frac{b-x}{b-a} & , a \leq x < b \\ 0 & , x \geq b \end{cases} \tag{6.8}$$

Equation for the MODERATE region:

$$Y = \begin{cases} 0 & , x < c \text{ and } x > e \\ \frac{x-c}{d-c} & , c \leq x \leq d \\ \frac{e-x}{e-d} & , d < x \leq e \end{cases} \tag{6.9}$$

Equation for the HIGH region:

$$Y = \begin{cases} 0 & , 0 < x < f \\ \frac{x-f}{g-f} & , f \leq x < g \\ 1 & , x \geq g \end{cases} \tag{6.10}$$

The values of variables from a to g are as follows:

Table 6.6: **Value of Graph Parameters for Fuzzy Variables**

	a	b	c	d	e	f	g
Number of hops	1	3	2	4	6	5	7
Minimum Energy	0.0	3.0	2.0	5.0	8.0	7.0	9.0
Pheromone Update	0.0	0.35	0.25	0.5	0.75	0.65	1.0

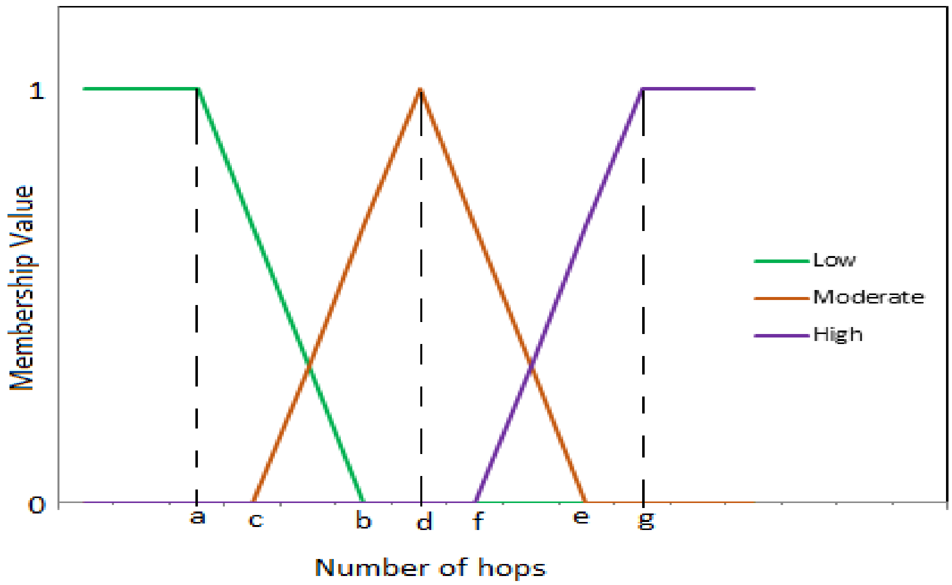

Fig. 6.2: **Membership function for Number of hops**

A path with lesser number of hops tends to increase the duration between the time the source sends a packet and the destination receives it. Therefore, other factors remain same, paths with smaller hop count should be preferred to improve the performance of the routing protocol.

In Fig. 6.2, the blue region represents the lower values of number of hops. A link with lower number of hops will have a higher value of pheromone update. The orange region represents the moderate values of number of hops. A link with moderate value of number of hops will have a moderate value of pheromone update.

The grey region represents the higher values of number of hops. A link with high value of number of hops will have a low value of pheromone update.

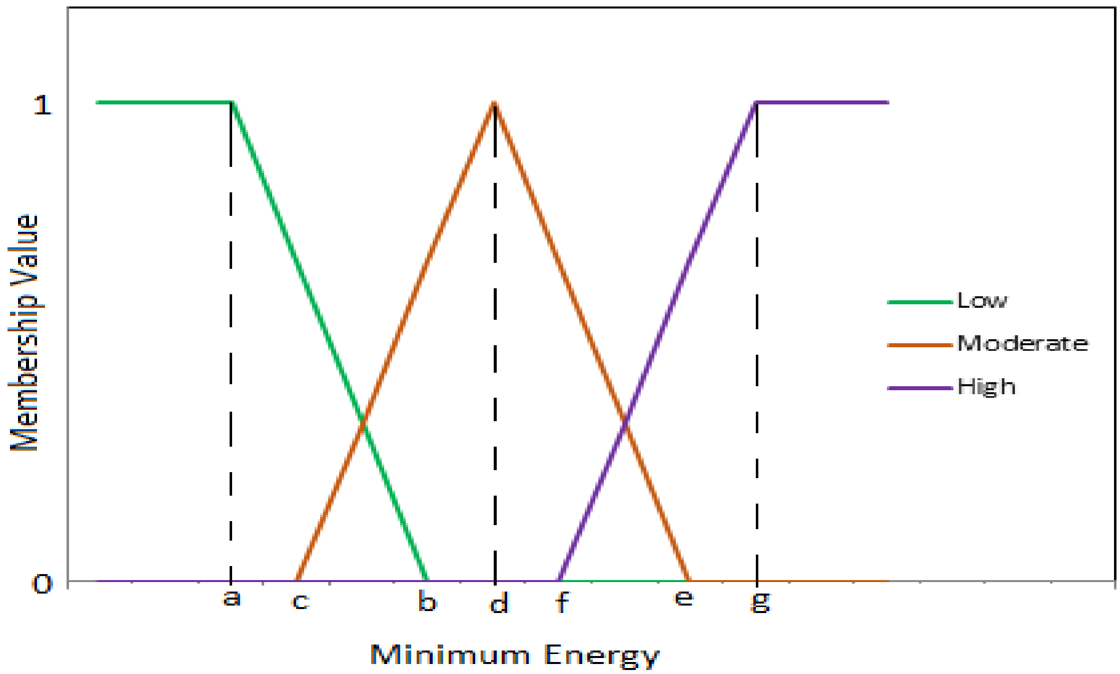

Fig. 6.3: Membership function for minimum energy

A path with greater value of minimum energy of intermediate nodes present in the path is less likely to go down earlier than a path with a smaller value for the minimum, energy of intermediate nodes along it. Therefore, other factors remaining same paths with greater value of minimum energy of intermediate nodes present in the path should be preferred to improve the performance of the routing protocol.

In Fig. 6.3, the blue region represents the lower values of minimum energy. A link with lower value of minimum energy will have a lower value of pheromone update due to low link stability. The orange region represents the moderate values of minimum energy. A link with moderate value of minimum energy will have a moderate value of pheromone update. The grey region represents the higher values of minimum energy. A link with high value of minimum energy will have a higher value of pheromone update.

The rules to calculate the value of pheromone update are as follows:

1. If number of hops are *high* AND minimum energy is *low*

   THEN pheromone update is *low*.

   OR

2. If number of hops are *moderate* AND minimum energy is *moderate*

   THEN pheromone update is *moderate*.

   OR

3. If number of hops are *low* AND minimum energy is *high*

   THEN pheromone update is *high*.

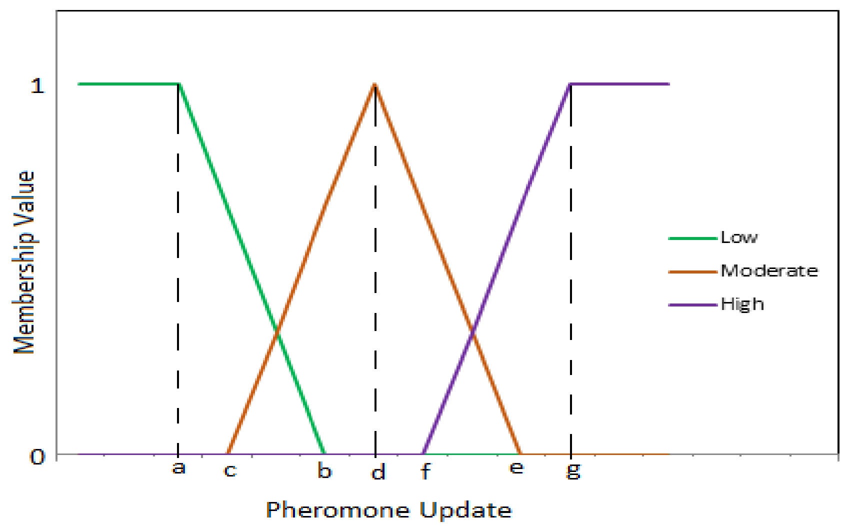

Fig. 6.4: Membership function for Pheromone Update

6.3.7 Proactive Forward Ant (PRFA)

Each node, which is the source of a communication session periodically, sends proactive forward ants in search of better paths than the current best path. The destinations for which

PRFAs are to be sent is selected randomly from the active destination table, D^i of a node i. Proactive forward ants are unicasted towards the destination, using the same probabilistic routing rule as reactive forward ants.

Table 6.7: **Fields in a PRFA Packet**

S. No.	Field Name	Purpose
1.	Ant Direction	Specifies whether the ant is a forward or a backward ant. It is set to “Forward” for a PRFA packet
2.	Ant Type	Specifies whether the ant is a reactive or proactive ant packet. It is set to “Proactive” for a PRFA
3.	Sequence Number	Specifies sequence number of the PRFA packet
4.	Source	Specifies the address of the node which originated the PRFA packet
5.	Destination	Specifies the address of the node to which the source node wants to discover a path
6.	Memory Size	Specifies number of entries in the memory of the PRFA packet
7.	Memory	It is a list which stores address of nodes visited by the PRFA packet

Unlike reactive ants, proactive ants are never broadcasted, and when they arrive at a node where there is no pheromone available they are just discarded. The main function of these proactive forward ants is to find paths to the destination other than the path currently being used so that if the link goes down or path becomes non-existent routing can still take place without the need of sending reactive ants from the source node. The fields present in a PRFA packet are summarized in Table 6.7.

6.3.8 Proactive Backward Ant (PRBA)

When a proactive ant arrives at the destination, a backward ant identical to the ones used during the reactive path setup is sent back to the source. This way, promising paths other than the current path can be found. This increases the number of paths available for data routing, which grows to a full mesh, and allows the algorithm to exploit new routing opportunities in the ever-changing topology. The rules for updating pheromone table by PRBA packets are the same as those used for updating pheromone table by REBA packets. The fields present in a PRBA packet are summarized in the following table:

Table 6.8: **Fields in a PRBA Packet**

S. No.	Field Name	Purpose
1.	Ant Direction	Specifies whether the ant is a forward or a backward ant. It is set to “Backward” for a PRBA packet
2.	Ant Type	Specifies whether the ant is a reactive or proactive ant packet. It is set to “Proactive” for a PRBA
3.	Source	Specifies the address of the node which originated the PRFA packet corresponding to the PRBA packet
4.	Destination	Specifies the address of the node to which the source node wants to discover a path
5.	Memory Size	Specifies number of entries in the memory of the PRBA packet
6.	Memory	It is a list which stores address of nodes which need to be visited by the PRBA packet to backtrack to the source node
7.	Hop Count	Specifies the number of hops visited by the PRBA packet.
8.	Minimum Energy	Specifies the energy of the node having the minimum energy among the nodes visited by the PRBA packet

6.3.9 Procedural description of the protocol

6.3.9.1 Procedure ForwardDataPacket: Called when source node s needs to send data packets to destination node d

1. if (routing information for d exists) then

Randomly generate next hop from pheromone table according to pheromone

values and forward packet

else

Create new REFA

REFA.source ← s

REFA.destination ← d

REFA.sequenceNumber ← counter++

REFA.memory[0] ← s;

REFA.memsize ← 1;

Broadcast REFA

6.3.9.2 Procedure ReceiveREFA: Called when a node i receives a REFA

1. s ←REFA.source
2. d ←REFA.destination
3. seq_num = REFA.sequenceNumber
4. if (i!=d ^ C^{i}_{sd} ≥ seq_num) then

   drop REFA and return
5. C^{i}_{sd} ← seq_num

6. REFA.memory[REFA.memsize++] ← i;
7. if (the node i is not the destination d) then

   if (routing information for d exists) then

   Randomly generate next hop from pheromone table according to pheromone values and forward packet

   else

   Broadcast REFA

   else

   Convert REFA to REBA and set following initial values:

   REBA.hop_count ← 0

   REBA.min_energy ← ∞

   Unicast REBA backwards to node REBA.memory[REBA.memsize-2]

6.3.9.3 Procedure Receive REBA: Called when a node i receives a REBA

1. s ← REBA.source
2. d ← REBA.destination
3. n ← REBA.memory[REBA.memsize-1]
4. REBA.memsize ← REBA.memsize -1
5. REBA.hop_count ← REBA.hop_count + 1
6. REBA.min_energy ← min(REBA.min_energy, energy of current node)
7. ${N^d}_i$ ← REBA.hop_count
8. ${E^d}_i$ ← REBA.min_energy

9. Compute pheromone update, $\tau_{nd}^{i} = Fuzzy(N_i^d, E_i^d)$

10. Update pheromone table, $\mathrm{T}_{\mathrm{nd}}^{\mathrm{i}} \leftarrow \alpha * \mathrm{T}_{\mathrm{nd}}^{\mathrm{i}} + (1-\alpha) * \tau_{\mathrm{nd}}^{\mathrm{i}}$

11. if (the node i is same as source node s) then

    drop REBA

    else

    Unicast REBA backwards to node
    REBA.memory[REBA.memsize-2]

6.3.9.4 Procedure Send PRFA: Called when a node i needs to send PRFA

1. if (Active Destination Table of node i, D^i is empty) then
   return

   else

   d ← randomly select destination from D^i

2. PRFA.source ← i

3. PRFA.destination ← d

4. PRFA.sequenceNumber ← counter++

5. PRFA.memory[0] ← s;

6. PRFA.memsize ←1;

7. if (routing information for d exists) then

   Randomly generate next hop from pheromone table according to

   pheromone values and forward PRFA

   else

   Broadcast PRFA

6.3.9.5 Procedure Receive PRFA: Called when a node i receives a PRFA

1. s ←PRFA.source
2. d ← PRFA.destination
3. seq_num = REFA.sequenceNumber
4. if (${C^i}_{sd}$ ≥ seq_num) then

   drop PRFA and return
5. ${C^i}_{sd}$ ← seq_num
6. PRFA.memory[PRFA.memsize++] ← i;
7. if (the node i is not the destination) then

   if (routing information for d exists) then

   Randomly generate next hop from pheromone table according to

   pheromone values

   and forward packet

   else

   drop PRFA

   else

   Convert PRFA to PRBA and set following initial values:

   PRBA.hop_count ← 0

   PRBA.min_energy ←∞

   Unicast PRBA b

6.3.9.6 Procedure Receive PRBA: Called when a node i receives a PRBA

1. s ←PRBA. source
2. d ←PRBA.destination
3. n ← PRBA.memory[PRBA.memsize-1]

4. PRBA.memsize ← PRBA.memsize -1
5. PRBA.hop_count ← PRBA.hop_count + 1
6. PRBA.min_energy ← min(PRBA.min_energy, energy of current node)
7. ${N^d}_i$ ← PRBA.hop_count
8. ${E^d}_i$ ← PRBA.min_energy
9. Compute pheromone update, $\tau_{nd}^{i} = Fuzzy(N_i^d, E_i^d)$
10. Update pheromone table, $\mathrm{T}_{\mathrm{nd}}^{\mathrm{i}} \leftarrow \alpha * \mathrm{T}_{\mathrm{nd}}^{\mathrm{i}} + (1-\alpha) * \tau_{\mathrm{nd}}^{\mathrm{i}}$
11. if (the node i is same as source node s) then

    drop PRBA

    else

    Unicast PRBA backwards to node PRBA.memory[PRBA.memsize-2]

6.4 Simulation Results and Analysis

The Performance of proposed protocol is evaluated by Network Simulator 2(NS2) .In this simulation, we consider a network with 10, 20, 30,40,50,60,70,80,90, and 100 sensor nodes that are randomly moves in a dimension of 500 meter × 500 meter area for 100 seconds. The simulation settings and parameters are as in summarized as per Table 6.9.

Table 6.9: Simulation Parameters

Parameter	Simulation Value
Antenna type	Omni Directional Antenna
Channel Type	Wireless channel
Radio-propagation model	Two Ray Ground
MAC type	IEEE 802.11

Simulation area	500*500 m^2
Node velocity	5ms^{-1} to 10 ms^{-1}
Node movement model	Random Waypoint
Traffic type	CBR
Packet Size	512 bytes
Pause time	5 s
Simulation time	100 s

6.4.1 Evaluation Metrics

The performance evaluation involves the following metrics:

- Average End to End Delay: This parameter represents the average time for transmitting a data packet from source to destination.
- Normalised Routing Load: This parameter tells us about the quantity and size of control packets generated by the protocol to discover and hold a route. It is defined as the ratio of the number of generated control packets to the number of generated data packets.
- Packet Delivery Ratio: It is the rate of packets successfully delivered.
- Average Throughput: It is the rate of successful message delivery over a communication channel.

Fig. 6.5 below depicts that packet delivery fraction in proposed protocol is more than in standard AODV. This behaviour is also credited to multipath routing since the availability of multiple paths allows an alternative route to be followed to a destination in case another one goes down due to a node or link failure without requiring the initiation of another route discovery. This results in more packets being delivered as compared with AODV in which each node which maintains only a single path to the destination node.

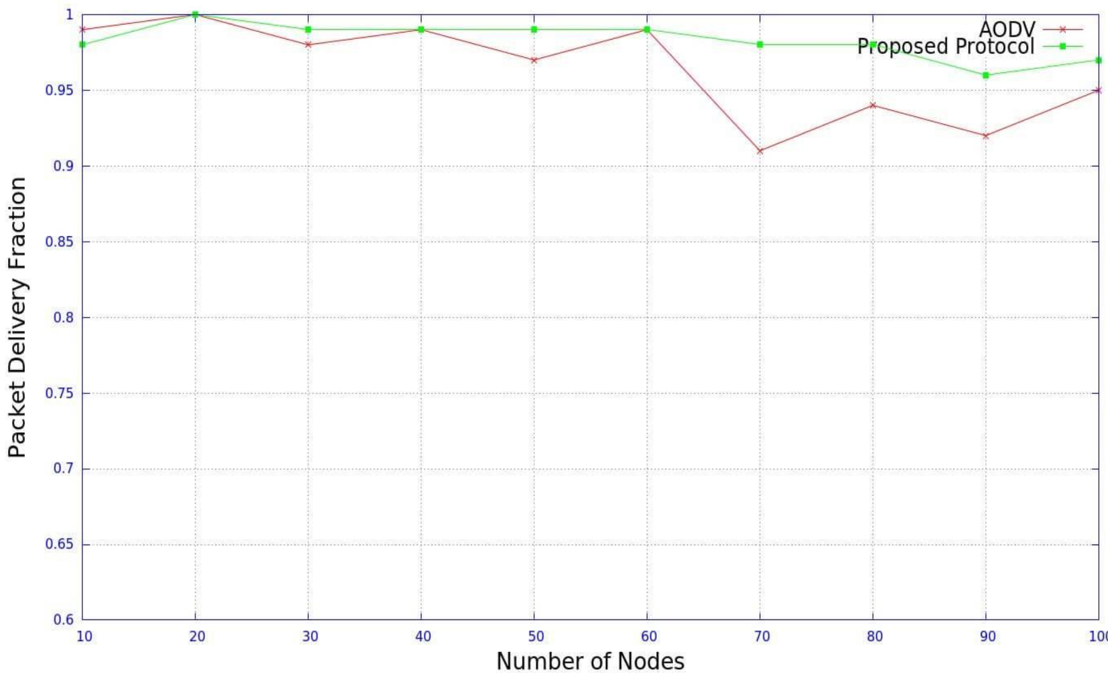

Fig. 6.5: Packet Delivery Fraction vs Number of nodes

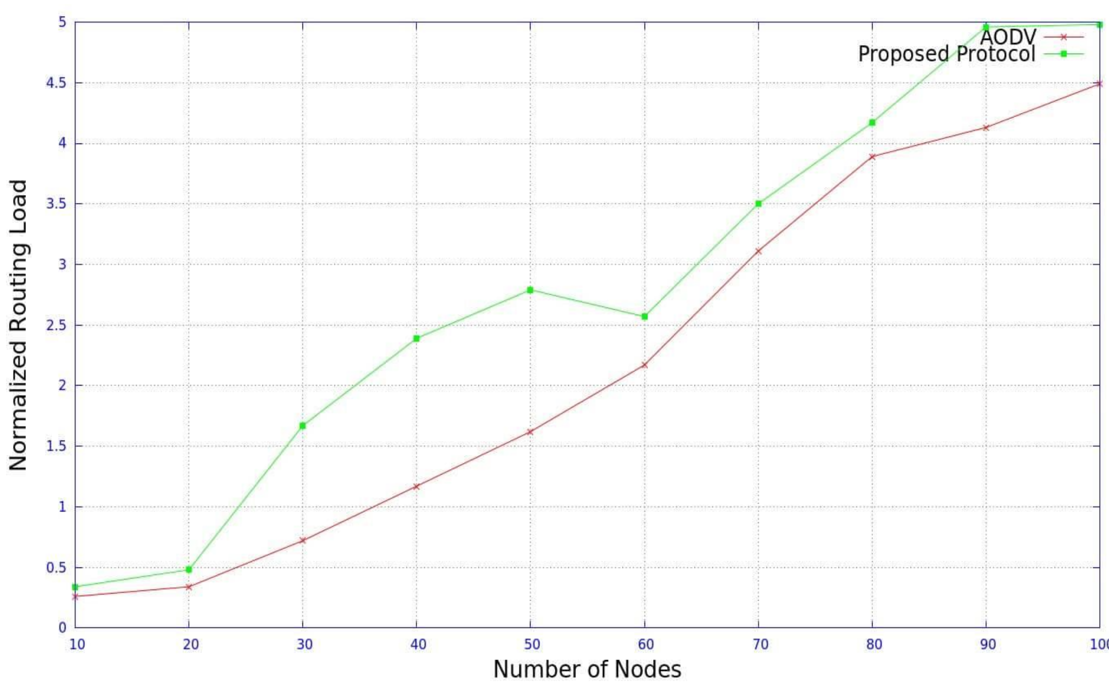

Fig. 6.6: Normalized Routing Load vs Number of Nodes

The Fig. 6.6 above infers that the normalized routing load in proposed protocol tends to be greater than that for AODV. This is attributed to the hybrid nature of proposed protocol. Being both reactive and pro-active in nature, it tends to send out “Proactive Ants” to maintain existing paths and discover new ones. Sending out Proactive routing packets increases the number of routing packets in the network and hence results in a much greater value of normalised routing load for our proposed protocol. In contrast, AODV is a purely reactive routing protocol and hence sends out routing packets only when a source node needs to send data packets to a destination node but the source node does not have any information stored for the destination nodes in its routing table. Although this type of reactive behaviour incurs more time to discover paths ,but it effectively lowers the routing overhead

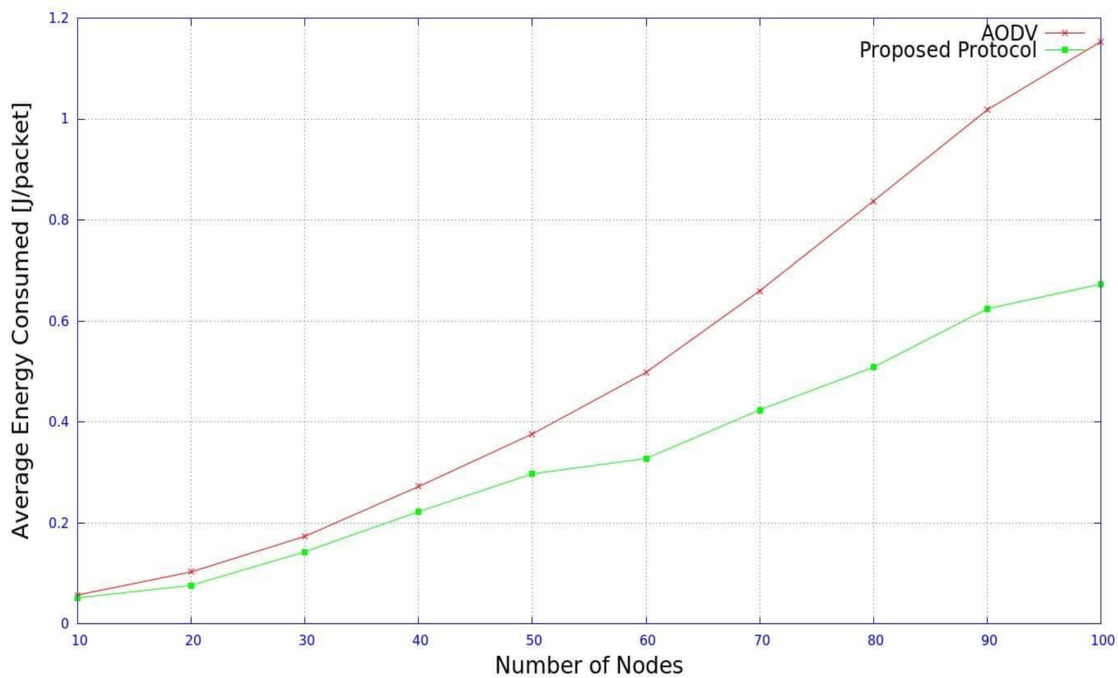

Fig. 6.7: Average Energy Consumption vs Number of Nodes

The Fig. 6.7 depicts that average energy consumed per data packet in proposed protocol is less than the standard AODV protocol. This is due to the fact that multipath routing tends to distribute the load in the network among multiple paths and hence effectively eliminates bottlenecks. Also in the case of AODV, bottleneck nodes i.e. nodes which are responsible for routing a high amount of network traffic tend to loose energy quickly and may become “dead” quite early rendering them unavailable to participate in the routing of packets in the network.

The Fig. 6.8 depicts that the average throughput in our proposed protocol is more than in standard AODV. This behaviour is also credited to multipath routing since the availability of multiple paths allows an alternative route to be followed to a destination in case another one goes down due to a node or link failure without requiring the initiation of another route discovery.

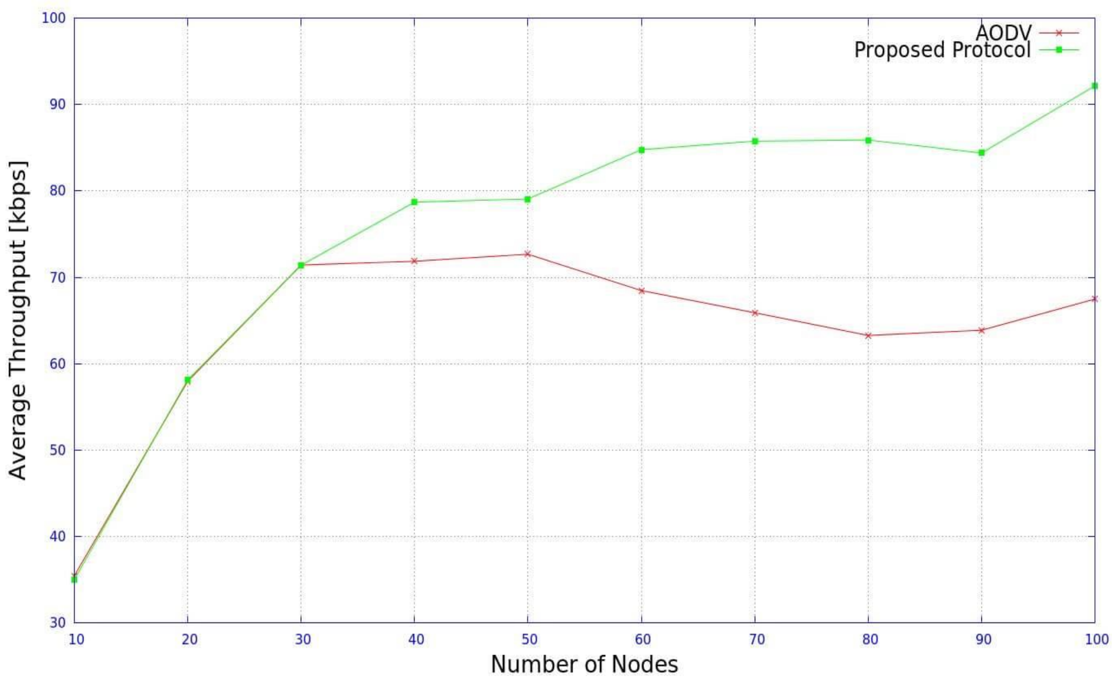

Fig. 6.8: Average Throughput vs Number of Nodes

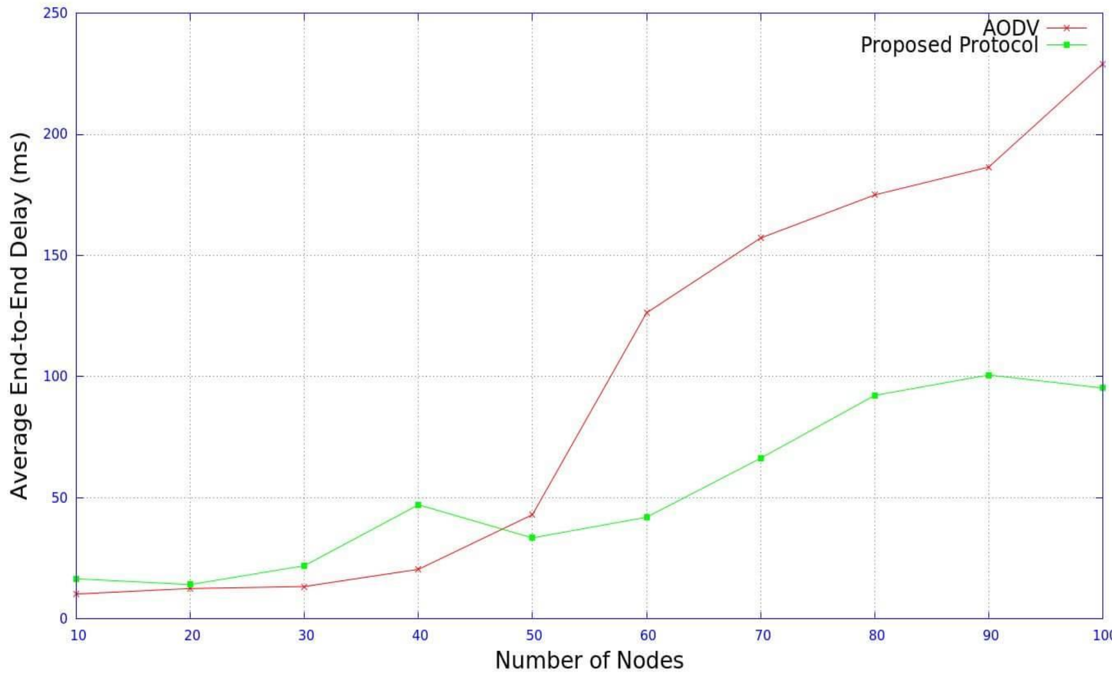

Fig. 6.9: Average end to end delay vs Number of Nodes

The average end-to-end (ETE) delay refers to the average time taken by a data packet to be transmitted across a network from the source to destination (Sink). Since it is desirable to deliver data packets to the destination, nodes with as much less delay as possible, a routing protocol must have a comparatively lower value for average end-to-end delay (ETE). In cases when there are more path failures, the proposed protocol may use another established route to send packets whereas a route discovery phase has to be initiated in AODV, which results in more delay due to additional time required to discover routes. Hence, multipath routing presents a trade-off for end-to-end delay. From the Fig. 6.9 above we can infer that the proposed protocol performs considerably well compared to AODV with respect to end-to-end delay, indicating that path failures tend to dominate over longer paths. Also as number of nodes increases, Route Request packets in AODV broadcasted to more number of nodes increasing the load on the network and affecting its performance largely.

6.5 Conclusion

The proposed approach is based on the principles of swarm intelligence, which are inherent in ant colony optimization. The goodness of a path is evaluated by a set of fuzzy rules, which consider both the number of hops in it and the energy of the node having minimum energy among nodes in the path. The performance of the proposed approach is compared to a standard state of the art routing protocol, namely AODV (Ad-hoc on demand distance vector routing).

- The proposed approach performs better than the standard AODV protocol in terms of delivering packets and energy efficiency at the cost of using proactive routing packets to search for new paths and maintain existing ones.
- The result depicts that average energy consumed per data packet in proposed protocol is less than the standard one. This is due to the fact that multipath routing tends to distribute the load in the network among multiple paths and hence effectively eliminates bottlenecks.
- The results also depicts that the average throughput in proposed protocol is more than in standard. This behaviour is also credited to multipath routing since the availability of multiple paths allows an alternative route to be followed to a destination in case another one goes down due to a node or link failure without requiring the initiation of another route discovery.

CHAPTER 7

CONCLUSIONS AND FUTURE SCOPE

In this research, we have developed four new algorithms for Data Aggregation & Fusion in the wireless sensor network. First algorithm is based on the probabilistic clustering approach, which maintains the requirement of clusters of almost same size and the optimal percentage of CHs in each round and also used differential temporal and energy based aggregation to optimize lifetime of the wireless sensor network by using optimal differential inter cluster head election latency determination and optimal differential energy based cluster head selection approach. The second algorithm uses particle swarm optimization (PSO) based approach with discrete search space, the success of our proposed algorithm depends greatly on the formulation of the fitness function. Therefore, we have defined a fitness function that includes all optimization criteria. The third algorithm is based on Grid-based coverage calculation strategy and uses dynamic positioning of a mobile sensors using Artificial Bee Colony Algorithm. In the fourth proposed algorithm, we have explored routing approach, which exploits the principles of swarm intelligence inherent in ant colony optimization (ACO) for transmission of data in which goodness of a path is evaluated by a set of fuzzy rules. The exhaustive experimentation and analysis show that the proposed algorithms improve the degree of optimization in the performance of sensor network and enhances the lifetime of wireless sensor networks.

There are many open issues related to the research in this thesis. Some possible future research directions are listed below

- Specific aggregation function is not considered for example duplicate sensitive may be investigated.
- Analytical approach may be investigated to distributed multi sensor data aggregation & fusion problems where there are time varying communication bandwidth constraints
- Spatial and temporal correlation among the sensor nodes based on coverage and location may be investigated.
- Coverage preservation by node scheduling may be investigated.

REFERENCES

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, "A survey on sensor networks," *Communications Magazine, IEEE,* vol. 40, no. 8, pp. 102- 114, 2002.

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

[3] M. Hamdi, N. Boudriga and M. S. Obaidat, "Who moves: An optimized broadband sensor network for military vehicle tracking," *Int'l Journal of Communication Systems,* vol. 21, no. 3, pp. 277- 300, 2008.

[4] G. Simon, M. Maroti, A. Ledeczi, G. Balogh, B. Kusy, A. Nadas, G. Pap, J. Sallai and K. Frampton, "Sensor network based countersniper system," in *Proc. 2nd ACM Conf. on Embedded Networked Sensor Systems (SenSys '04)*, Baltimore, MD,pp. 1–12, 2004.

[5] E. A. Basha, S. Ravela and D. Rus, "Model-based monitoring for early warning flood detection," in *Proc. 6th ACM Conf. on Embedded Network Sensor Systems(SenSys '08)*, Raleigh, NC, pp. 295 - 308, 2008.

[6] A. Cerpa, J. Elson, M. Hamilton, J. Zhao, D. Estrin and L. Girod, "Habitat monitoring: application driver for wireless communications technology," in *Proc. Workshop on Data Communication in Latin America and the Caribbean (SIGCOMM LA '01)*, San Jose, Costa Rica, pp. 20 - 41, 2001.

[7] W. Hu, N. Bulusu, C. T. Chou, S. Jha, A. Taylor and V. N. Tran, "Design and evaluation of a hybrid sensor network for cane toad monitoring," *ACM Transactions on Sensor Networks (TOSN),* vol. 5, no. 1, pp. 1-28, 2009.

[8] J. Kim, J. Lim, J. Friedman, U. Lee, M. Srivastava, M. Gerla and D. Rosso, "Sewersnort: A drifting sensor for in-situ sewer gas monitoring," in *Proc. 6th Annual IEEE Communications Society Conf. on Sensor, Mesh and Ad Hoc Communications and Networks (SECON 09)*, Rome, Italy, 2009.

[9] A. Mainwaring, D. Culler, J. Polastre, R. Szewczyk and J. Anderson, "Wireless sensor networks for habitat monitoring," in *Proc. 1st ACM Int'l Workshop on Wireless Sensor Networks and Applications (WSNA '02)*, Atlanta, GA, pp. 88-97, 2002.

[10] M. V. Ramesh., "Real-time wireless sensor network for landslide detection," in *Proc. 2009 3rd Int'l Conf. on Sensor Technologies and Applications (SENSORCOMM '09)*, Athens/Glyfada, Greece, pp. 158 - 163, 2009.

[11] P. Sikka, P. Corke, P. Valencia, C. Crossman, D. Swain and G. Bishop-Hurley, "Wireless adhoc sensor and actuator networks on the farm," in *Proc. 5th Int'l Conf. on Information Processing in Sensor Networks(IPSN '06)*, Nashville, TN, pp. 492 - 499, 2006.

[12] R. Szewczyk, A. Mainwaring, J. Polastre, J. Anderson and D. Culler, "An analysis of a large scale habitat monitoring application," in *Proc. 2nd ACM Conf. on Embedded Networked Sensor Systems (SenSys '04)*, Baltimore, MD, pp. 214–226, 2004.

[13] F. Han, Y. Chang, P. Fan and A. V. Vasilakos, "Collaborative beamforming and dirty paper coding assisted transmission strategy for e-health wireless sensor networks," in *Proc. 5th Int'l Wireless Communications and Mobile Computing Conf(IWCMC '09)*, Leipzig, Germany, pp. 1080 -1084, 2009.

[14] H. Huo, Y. Xu, H. Yan, S. Mubeen and H. Zhang, "An elderly health care system using wireless sensor networks at home," in *Proc. 3rd Int'l Conf. on Sensor Technologies and Applications (SENSORCOMM '09)*, Athens/Glyfada, Greece, pp. 158 - 163, 2009.

[15] L. Schwiebert, S. K. Gupta and J. Weinmann, "Research challenges in wireless networks of biomedical sensors," in *Proc. 7th Annual Int'l Conf. on Mobile Computing and Networking (MobiCom '01)*, Rome, Italy, pp. 151 - 165, 2001.

[16] S. Kim, S. Pakzad, D. Culler, J. Demmel, G. Fenves, S. Glaser and M. Turon, "Wireless sensor networks for structural health monitoring," in *Proc. 4th ACM Conf. on Embedded Networked Sensor Systems (SenSys '06)*, Boulder, CO, pp. 427 - 428, 2006.

[17] N. Xu, S. Rangwala, K. K. Chintalapudi, D. Ganesan, A. Broad, R. Govindan and D. Estrin, "A wireless sensor network for structural monitoring," in *Proc. 2nd ACM Conf. on Embedded Networked Sensor Systems (SenSys '04)*, Baltimore, MD, pp. 13-24, 2004.

[18] J. Fogarty, C. Au and S. E. Hudson, "Sensing from the basement:a feasibility study of unobtrusive and low-cost home activity recognition," in *Proc.19th Annual ACM Symposium on User Interface Software and Technology(UIST '06)*, Montreux, Switzerland, pp. 91 - 100, 2006.

[19] Y. Kim, T. Schmid, Z. M. Charbiwala, J. Friedman and M. B. Srivastava, "NAWMS: nonintrusive autonomous water monitoring system," in *Proc. 6th ACM*

Conf. on Embedded Network Sensor Systems (SenSys ’08), Raleigh, NC, pp. 309–322, 2008.

[20] I. Stoianov, L. Nachman, S. Madden and T. Tokmouline, "Pipeneta wireless sensor network for pipeline monitoring," in *Proc. 6th Int’l Conf. on Information Processing in Sensor Networks (IPSN ’07)*, Cambridge, MA, pp. 264 - 273, 2007.

[21] C. Alippi, R. Camplani, C. Galperti and M. A. r. a. s.-p. W. f. f. a. e. m. Roveri, " A robust, adaptive, solar-powered WSN framework for aquatic environmental monitoring," *IEEE Sensors journal,* vol. 11, no. 1, pp. 45 - 55, 2011.

[22] N. Suryadevara, S. Mukhopadhyay, S. T. Kelly and S. S. Gill, "WSN-based smart sensors and actuator for power management in intelligent buildings," *IEEE/ASME Trans. Mecatronics,* vol. 20, no. 2, pp. 564 - 571, 2015.

[23] L. Chitnis, A. Dobra and S. Ranka, "Aggregation Methods for Large-Scale Sensor Networks," *ACM Transactions on Sensor Networks,* vol. 4, no. 2, pp. 1-36, 2008.

[24] S. Park, "Performance Evaluation of Data Aggregation Schemes in Wireless Sensor Networks," in *WCNC 2006 proceedings, IEEE Communications Society*, Las Vegas, NV, pp. 413 - 418, 2006.

[25] L. Yong-Min, W. Shu-Ci and N. Xiao-Hong, "The Architecture and Characteristics of Wireless Sensor Network," in *IEEE International Conference on Computer Technology and Development*, Kota Kinabalu, pp. 561-565, 2009.

[26] Y. Xue, Y. Cui and K. Nahrstedt, "Maximizing Lifetime for Data Aggregation in Wireless Sensor Networks," *Mobile Networks and Applications,* vol. 10, pp. 853 - 864, 2005.

[27] M. A. Sharaf, J. Beaver, A. Labrinidis, R. Labrinidis and P. K. Chrysanthis, "Balancing energy efficiency and quality of aggregate data in sensor networks," *VLDB J.,* vol. 13, no. 4, pp. 384-403, 2004.

[28] L. Galluccio, S. Palazzo and A. T. Campbell, "Efficient data aggregation in wireless sensor networks: an entropy-driven analysis," in *Proceedings of the IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '08),* Cannes, France, pp. 1-6, 2008.

[29] B. Singh and D. K. Lobiyal, "A novel energy-aware cluster head selection based on particle swarm optimization for wireless sensor networks," *Human-Centric Computing and Information Sciences,* vol. 2, no. 1, pp. 1-18, 2012.

[30] A. Sinha and D. K. Lobiyal, "Probabilistic Data Aggregation in Information based Clustered Sensor Network," *Wireless Personal Communications,* vol. 77, no. 2, pp. 1287-1310, 2014.

[31] Y. Ma, Y. Guo, X. Tian and M. Ghanem, "Distributed Clustering-Based aggregation Algorithm for Spatial Correlated Sensor Networks," *IEEE Sensors Journal*, vol. 11, no. 3, pp. 641-648, 2011.

[32] K.-W. Fan, S. Liu and P. Sinha, "On the potential of structure-free data aggregation in sensor networks," in *Proc. 25th Annual Joint Conf. of the IEEE Computer and Communications Societies (INFOCOM '06)*, Barcelona, Spain, pp. 1 - 12, 2006.

[33] F. Zhao, J. Shin and J. Reich, "Information-driven dynamic sensor collaboration," *IEEE Signal Processing Magazine*, vol. 19, no. 2, pp. 61-72, 2002.

[34] W. R. Heinzelman, A. Chandrakasan and H. Balakrishnan, "Energy efficient communication protocol for wireless micro sensor networks," in *Proceedings of the 33rd Hawaaian Interantional Conference on System Sciences*, pp. 3005 - 3014, 2000.

[35] M. C. Vuran and I. F. Akyildiz, "Spatial Correlation-based Collaborative Medium Access Control in Wireless Sensor Networks," *IEEE/ACM Trans. Networking*, vol. 14, no. 2, pp. 316 - 329, 2006.

[36] B. Lee and K. Lim, "An Energy-Efficient Hybrid Data-Gathering Protocol Based on the Dynamic Switching of Reporting Schemes in Wireless Sensor Networks," *IEEE Systems Journal*, vol. 6, no. 3, pp. 378-387, 2012.

[37] W. B. Heinzelman, A. P. Chandrakasan and H. Balakrishnan, "An application-specific protocol architecture for wireless micro sensor networks," *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, pp. 660-70, 2002.

[38] B. Krishnamachari, D. Estrin and S. B. Wicker, "The impact of data aggregation in wireless sensor networks," in *ICDCSW '02: Proceedings of the 22nd International Conference on Distributed Computing Systems.* , Washington, DC, USA: IEEE Computer Society, 2002.

[39] T. Abdelzaher, T. He and J. Stankovic, "Feedback control of data aggregation in sensor networks," in *Proc. 43rd IEEE Conf. on Decision and Control*, Atlantis, Paradise Island, Bahamas, pp.1490 -1495, 2004.

[40] S. Gobriel, S. Khattab, D. Mosse, J. Brustoloni and R. Melhem, "Ridesharing: Fault tolerant aggregation in sensor networks using corrective actions," in *Proc. 3rd Annual Conf. on Sensor, Mesh and Ad Hoc Communications and Networks (SECON '06)*, Reston, VA, pp. 595 - 604, 2006.

[41] S. Madden, M. J. Franklin, J. M. Hellerstein and W. Hong, "TAG: a tiny aggregation service for ad-hoc sensor networks," *SIGOPS Oper. Syst. Rev*, vol. 36(SI), pp. 131-146, 2002.

[42] I. Solis and K. Obraczka, "The impact of timing in data aggregation for sensor networks," in *Proc. 2004 Int'l Conf. on Communications (ICC '04)*, Paris, France, pp. 3640–3645, 2004.

[43] H. Zhou, Y. Wu, Y. Hu and G. Xie, "A novel stable selection and reliable transmission protocol for clustered heterogeneous Wireless sensor networks," *Computer Communications,* vol. 33, no. 15, pp. 1843 - 1849, 2010.

[44] J. Gao, L. Guibas, N. Milosavljevic and J. Hershberger, "Sparse data aggregation in sensor networks," in *Proc. 6th Int'l Conf. on Information Processing in SensorNetworks (IPSN '07)*, Cambridge, MA, pp. 430- 439, 2007.

[45] T. He, B. M. Blum, J. A. Stankovic and T. Abdelzaher, "AIDA:Adaptive application-independent data aggregation in wireless sensor networks," *ACM Transactions on Embedded Computing Systems,* vol. 3, no. 2, pp. 426 - 457, 2004.

[46] C. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann and F. Silva, "Directed diffusion for wireless sensor networking," *IEEE/ACM Transactions on Networking,* vol. 11, pp. 2-16, 2003.

[47] J. Kulik, W. Heinzelman and H. Balakrishna, "Negotiation-based protocols for disseminating information in wireless sensor network," in *Proc. 5th Annual Int'l Conf. on Mobile Computing and Networking,* Seattle, WA, pp. 169 - 185, 2002.

[48] R. Beraldi, R. Baldoni and R. Prakash, "Lukewarm potato forwarding: A biased random walk routing protocol for wireless sensor networks," in *Proc. 6th Annual IEEE Communications Society Conf. on Sensor, Mesh and Ad Hoc Communications and Networks*, Rome, Italy, 2009.

[49] E. Felemban, C.-G. Lee, E. Ekici, R. Boder and S. Vural, "Probabilistic QoS guarantee in reliability and timeliness domains in wireless sensor networks," in *Proc. 24th Int'l Annual Joint Conf. of the IEEE Compute and Communications Societies(INFOCOM '05)*, Miami, FL, pp. 2646 - 2657, 2005.

[50] S. Ganeriwal, R. Kumar and M. B. Srivastava, "Timing-sync protocol for sensor networks," in *Proc. 1st ACM Conf. on Embedded Networked Sensor Systems (SenSys '03)*, Los Angeles, CA, pp. 138- 149, 2003.

[51] J. Koo, R. K. Panta, S. Bagchi and L. Montestruque, "A tale of two synchronizing clocks," in *Proc. 7th ACM Conf. on Embedded Networked Sensor Systems (SenSys '09)*, Berkeley, CA, pp. 239 - 252, 2009.

[52] J. Feng, D. Eager and D. Makaroff, "Aggregation protocols for high rate, low delay data collection in sensor networks," in *Proc. 8th Int'l IFIP-TC 6 Networking Conf. (Networking '09)*, Aachen, Germany, pp. 26 - 39, 2009.

[53] J. Feng, D. Eager and D. Makaroff, "Asynchronous data aggregation for real-time monitoring in sensor networks," in *Proc. 6th Int'l IFIP-TC 6 Networking Conf. (Networking '07)*, Atlanta, GA, pp. 73 - 84, 2007.

[54] A. Sinha and D.K.Lobiyal, "A Multi-Level Strategy for Energy Efficient Data Aggregation in Wireless Sensor Networks," *Wireless Personal Communications*, vol. 72, no. 2, pp. 1513-1531, 2013.

[55] S. Motegi, K. Yoshihara and H. Horiuchi, "DAG based in-network aggregation for sensor network monitoring," in *Proc. Int'l Symposium on Applications on Internet (SAINT '06)*, Phoenix, AZ, pp. 292 - 299, 2006.

[56] W. Xiaoling, S. Lei, W. Jin, J. Cho and S. Lee, "Energy efficient deployment of mobile sensor network by PSO," in *Advanced Web and Network Technologies, and Applications: APWeb 2006 International Workshops: XRA, IWSN,MEGA, and ICSE*, Harbin, China, 2006.

[57] J. Kennedy and R. C. Eberhart, "Particle swarm optimization," in *Proceedings of IEEE International Conference on Neural Networks*, Perth, Australia, vol. 4, pp. 1942-1948, 1995.

[58] R. C. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *Proceedings of the Sixth International Symposium on Micromachine and Human Science*, Nagoya, Japan, pp. 39-43, 1995.

[59] D. Karaboga, "An Idea based on Honey Bee Swarm for Numerical Optimization," Tech. Rep. TR06, Erciyes University, Engineering Faculty, Computer Engineering Department, pp. 1 - 10, 2005.

[60] D. Karaboga and B. Akay, "A comparative study of Artificial Bee Colony algorithm," *Applied Mathematics and Computation*, vol. 214, pp. 108 -132, 2009.

[61] M. Dorigo, V. Maniezzo and A. Colorni, "Ant system: optimization by a colony of cooperating agents," *IEEE Transactions on Systems,Man, and Cybernetics, Part B (Cybernetics)*, vol. 26, no. 1, pp. 29 - 41, 1996.

[62] M. Dorigo and T. Stützle, "Ant colony optimization: overview and recent advances," in *Gendreau M, Potvin JY (eds) Handbook of metaheuristics. International series in operations research management science*, Vols. , vol 146, 2nd edn. Kluwer Academic, Dordrecht, pp , Dordrecht, Kluwer Academic, 2010, pp. 227 - 264.

[63] A. Colorni, M. Dorigo and V. Maniezzo, "Distributed Optimization by Ant Colonies," in *Proceeding of the 1st European Cnference on Artificial Life*, Paris, France,pp. 134 - 142, 1991.

[64] L. Zadeh, "Fuzzy sets," *Information and control,* vol. 8, no. 3, pp. 338 - 353, 1965.

[65] L. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning-I," *Information Sciences* , vol. 8, no. 3, pp. 199 - 249, 1975.

[66] H. Karl and A. Willig, Protocols and Architectures for Wireless Sensor Networks, John Wiley & Sons, 2005.

[67] G. J. Pottie and W. J. Kaiser, "Wireless integrated network sensors," *Communications of the ACM,* vol. 43, no. 5, pp. 51 - 58, 2000.

[68] T. He, J. A. Stankovic, C. Lu and T. Abdelzaher, "SPEED:A stateless protocol for real-time communication in sensor networks," in *Proc. 23rd Int'l Conf. on Distributed Computing Systems (ICDCS '03)*, Providence, RI, pp. 46–55, 2003.

[69] P. Skraba, Q. Fang, A. Nguyen and L. Guibas, "Sweeps over wireless sensor networks," in *Proc. 5th Int'l Conf. on Information Processing in Sensor Networks (IPSN '06)*, Nashville, TN, pp. 143-151, 2006.

[70] E. Fasoloy, M. Rossi and J. W. M. Zorzi, "In-network aggregation techniques for wireless sensor networks: a survey," *IEEE Wireless Communications,* vol. 14, no. 2, pp. 70-87, 2007.

[71] W. Zhang and G. Cao, "DCTC: dynamic convoy tree-based collaboration for target tracking in sensor networks," *IEEE transactions on wireless communications,* vol. 3, no. 5, pp. 1689 - 1701, 2004.

[72] W. Zhang and G. Cao, "Optimizing tree reconfiguration for mobile target tracking in sensor networks," in *Proc. 23rd Annual Joint Conf. of the IEEE Computer and Communications Societies (INFOCOM '04)*, Hongkong, China, pp. 2434 - 2445, 2004.

[73] K.-W. Fan, S. Liu and P. Sinha, "Scalable data aggregation for dynamic events in sensor networks," in *Proc. 4th ACM Conf. on Embedded Networked Sensor Systems (SenSys '06)*, Boulder, CO, pp. 181 -194, 2006.

[74] S. Bandyopadhyay and E. Coyle, "An energy efficient hierarchical clustering algorithm for wireless sensor networks," *Proc. INFOCOM 2003,* vol. 3, pp. 1713-1723, 2003.

[75] S. Nath, P. B. Gibbons, S. Seshan and Z. R. Anderson, "Synopsis diffusion for robust aggregation in sensor networks," in *Proc. 2nd ACM Conf. on Embedded Networked Sensor Systems (SenSys '04)*, Baltimore, MD, pp. 250-262, 2004.

[76] R. Bista, Y. K. Kim, Y. H. Choi and J. W. Chang, "A New Energy-Balanced Data Aggregation Scheme in Wireless Sensor Networks," in *2009 International*

Conference on Computational Science and Engineering, Vancouver, BC, pp. 558-563, 2009.

[77] C. Hui, S. Yufeng, W. Ye, L. T. Yang and X. Naixue, "Improvement of directed diffusion protocol based on data aggregation," in *2009 International Conference on Computational Science and Engineering*, Vancouver, BC, pp. 126-131, 2009.

[78] C. Intanagonwiwat, R. Govindan and D. Estrin, "Directed diffusion: a scalable and robust communication paradigm for sensor networks," in *Proc. 8th Annual Int'l Conf. on Mobile Computing and Networking (Mobi Com '00)*, Boston, MA, pp. 56 - 67, 2000.

[79] H. Weerasinghe, I. H. Elhajj, A. Krsteva and M. A. Najm, "Data centric adaptive innetwork aggregation for wireless sensor networks," in *IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, Switzerland, pp. 1-6, 2007.

[80] J. Mao, Z. Wua and X. Wuc, "A TDMA scheduling scheme for many to one communications in wireless sensor networks," *Computer Communications*, vol. 30, no. 4, pp. 863 - 872, 2007.

[81] V. Loscr, G. Morabito and S. Marano, "A two-levels hierarchy for low-energy adaptive clustering hierarchy (TL-LEACH)," in *62nd Vehicular Technology Conference (VTC' 2005fall)*, pp. 1809 - 1813, 2005.

[82] M. O. Farooq, A. B. Dogar and G. A. Shah, "MR-LEACH: Multi-hop routing with low energy adaptive clustering hierarchy," in *4th International Conference on Sensor Technologies and Applications (SENSORCOMM 2010)*, Venice/Mestre, Italy, pp. 262 - 268, 2010.

[83] A. Manjeshwar, Q.-A. Zeng and D. P. Agarwal, "An Analytical Model for Information Retrieval in Wireless Sensor Networks using Enhanced APTEEN Protocol," *IEEE Trans. Parallel and Distributed Systems*, vol. 13, no. 12, pp. 1290-1302, 2002.

[84] F. Xiangning and S. Yulin, "Improvement on leach protocol of wireless sensor network," in *International Conference on Sensor Technologies and Applications (SensorComm 2007)*, Valencia, Spain, pp. 260 - 264, 2007.

[85] M. B. Yassein, A. Al-zou'bi, Y. Khamayseh and W. Mardini, " Improvement on leach protocol of wireless sensor network (V-LEACH)," *Journal of Digital Content Technology and its Applications*, vol. 3, no. 2, pp. 132 - 136, 2009.

[86] G. Ran, H. Zhang and S. Gong, "Improving on LEACH protocol of wireless sensor networks using fuzzy logic," *Journal of Information & Computational Science*, vol. 7, no. 3, pp. 767-775, 2010.

[87] W. Jung, K. Lim, W. Ko and S. Park, "Efficient clustering-based data aggregation techniques for wireless sensor networks," *Wireless Netw,* vol. 17, pp. 1387 - 1400, 2011.

[88] O. Younis and S. Fahmy, "HEED: a hybrid, energy-efficient,distributed clustering approach for ad hoc sensor networks," *IEEE Transactions on Mobile Computing,* vol. 3, no. 4, pp. 366 - 379, 2004.

[89] S. Lindsey and C. Raghavendra, "PEGASIS: Power-efficient gathering in sensor information systems," in *Proceedings of IEEE Aerospace Conference*, vol. 3, pp. 1125 - 1130, 2002.

[90] V. Mhatre and C. Rosenberg, " Design guidelines for wireless sensor networks: communication, clustering and aggregation," *Ad Hoc Networks,* vol. 2, no. 1, pp. 45 - 63, 2004.

[91] Y. P. Chen, A. L. Liestman and J. Liu, "A hierarchical energy-efficient framework for data aggregation in wireless sensor networks," *IEEE Transactions on Vehicular Technology,* vol. 55, no. 3, pp. 789 - 796, 2006.

[92] S. Peng, S. Li, Y. Peng, W. s. Tang and N. Xiao, "Real-time data delivery in wireless sensor networks: A data-aggregated,cluster-based adaptive approach," in *Proc. 4th Int'l Conf. on Ubiquitous Intelligence and Computing (UIC '07)*, Hong Kong, China, pp. 514 - 523, 2007.

[93] K. Karenos, V. Kalogeraki and S. Krishnamurthy, "Cluster-based Congestion control for sensor network," *ACM Transaction on Sensor Networks,* vol. 4, pp. 1-39, 2008.

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

[95] K. Dilip, C. A. Trilok and R. B. Patel, "EEHC: energy efficient heterogeneous clustered scheme for wireless sensor networks," *Computer Communications,* vol. 32, no. 4, pp. 662 - 667, 2009.

[96] D. Wei, S. Kaplan and H.A. Chan, "Energy efficient clustering algorithms for wireless sensor networks," in *ICC Workshops - 2008 IEEE International Conference on Communications Workshops*, Beijing, pp. 236-240, 2008.

[97] Z. Beiranvand, A. Patooghy and M. Fazeli, "I-LEACH: An efficient routing Algorithm to Improve Performance & to Reduce Energy Consumption in Wireless Sensor Networks," in *IEEE 5th International Conference on Information and Knowledge Technology*, Shiraz, pp. 13-18, 2013.

[98] F. Yuan, Y. Zhan and Y. Wang, "Data Density Correlation Degree Clustering Method for Data Aggregation in WSN," *IEEE Sensors Journal,* vol. 14, no. 4, pp. 1089-1098, 2014.

[99] S. Chao, W. Ru-chuan, H. Hai-ping and S. Li-juan, "Energy efficient clustering algorithm for data aggregation in wireless sensor networks," *Journal of China Universities of Posts and Telecommunications,* vol. 17, no. 2, pp. 104 -109, 2010.

[100] A. Manjeshwar and D. P. Agrawal, "TEEN: A routing protocol for enhanced efficiency in wireless sensor networks," in *Proceedings of the 15th International Parallel and Distributed Processing Symposium,* San Francisco,CA,USA, pp. 2009-2015, 2001.

[101] A. Manjeshwar and D. P. Agrawal, "APTEEN: A hybrid protocol for efficient routing and comprehensive information retrieval in wireless sensor networks," in *16th International Parallel and Distributed Processing Symposium (IPDPS '02)*, Florida, USA, pp. 195- 202, 2002.

[102] J. Yuea, W. Zhang, W. Xiao, D. Tang and J. Tang, "Energy Efficient and Balanced Cluster-Based Data Aggregation Algorithm for Wireless Sensor Networks," in *International Workshop on Information and Electronics Engineering (IWIEE) , Procedia Engineering*, pp. 2009-2015, 2012.

[103] R. V. Kulkarni, A. Forster and G. K. Venayagamoorthy, "Computational intelligence in wireless sensor networks: A survey," *IEEE Communications Surveys & Tutorials,* vol. 13, no. 1, pp. 68-96, 2011.

[104] H. O. Tan and I. Korpeoglu, "Power efficient data gathering and aggregation in wireless sensor networks," *ACM SIGMOD Record,* vol. 32, no. 4, pp. 66-71, 2003.

[105] G. Smaragdakis, I. Matta and A. Bestavros, "SEP :A stable election protocol for clustered heterogeneous wireless sensor networks," in *Proceedings of the Second International Workshop on Sensor and Actor Network Protocols and Applications (SANPA 2004)*, Boston, MA, pp. 251-261,2004.

[106] M. Zhao, J. Li and Y. Yang, "A Framework of Joint Mobile Energy Replenishment and Data Gathering in Wireless Rechargeable Sensor Networks," *IEEE Transactions On Mobile Computing,* vol. 13, no. 12, pp. 2689-2705, 2014.

[107] T. Kang, J. Yun, H. Lee, I. Lee, H. Kim, B. Lee and K. Han, "A clustering method for energy efficient routing in wireless sensor networks," in *Proceedings of the 6th International Conference on Electronics, Hardware, Wireless and Optical Communications(WSEAS'07),* CorfuIceland,Greece, pp. 133 - 138, 2007.

[108] K. Vaidyanathan, S. Sur, S. Narravula and P. Sinha, "Data aggregation techniques in sensor networks," Technical report OSU-CISRC-11/04-TR60, Ohio State University, 2004.

[109] Y. Yao and J. Gehrke, "The cougar approach to in-network query processing in sensor networks," *ACM SIGMOD Record*, vol. 31, no. 3, pp. 9 - 18, 2002.

[110] M. Liliana, C. Arboleda and N. Nidal, "Comparison of clustering algorithms and protocols for wireless sensor networks," *in Proceedings of IEEE CCECE/CCGEI Conference*, pp. 1787 - 1792, 2006.

[111] N. M. A. Latiff, C. C. Tsimenidis and B. S. Sharif, "Energy aware clustering for wireless sensor networks using particle swarm optimization," in *Proceedings of the 18th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'07)*, Athens, Greece, pp. 1-5, 2007.

[112] A. A. Islam, C. S. Hyder, H. Kabir and M. Naznin, "Finding the Optimal Percentage of Cluster Heads from a New and Complete Mathematical Model on LEACH," *Wireless Sensor Network*, vol. 2, no. 2, pp. 129 - 140, 2010.

[113] S. Chatterjea and P. Havinga, "A dynamic data aggregation scheme for wireless sensor networks," in *ProRISC 2003, 14th Workshop on Circuits, Systems and Signal Processing*, Veldhoven,Netherlands, 2003.

[114] R. Rajagopalan and P. K. Varshney, "Data aggregation techniques in sensor networks: A survey," *IEEE Communications Surveys and Tutorials*, vol. 8, pp. 48-63, 2006.

[115] K. Dasgupta, K. Kalpakis and P. Namjoshi, "An efficient clustering-based heuristic for data gathering and aggregation in sensor networks," in *IEEE Wireless Communications and Networking, WCNC 2003*, New Orleans, LA, USA, pp. 1948 - 1953, 2003.

[116] G. Gupta and M. Younis, "Load-balanced clustering in wireless sensor networks," *in Proc. IEEE Int. Conf. Commun. ICC '03*, vol. 3, pp. 1848-1852, 2003.

[117] C. Vimalarani, R. Subramanian and S. N. Sivanandam, "An Enhanced PSO-Based Clustering Energy Optimization Algorithm for Wireless Sensor Network," *The Scientific World Journal*, vol. 2016, pp. 1-11, 2016.

[118] A. Gopakumar and L. jacob, "Localization in wireless sensor networks using particle swarm optimization," in *Proceedings of the IET International conference on Wireless,Mobile and Multimedia Networks*, Beijing ,China, pp. 227-230, 2008.

[119] K. Fan, S. Liu and P. Sinha, "Structure-Free Data Aggregation in Sensor Networks," *IEEE Transactions on Mobile computing*, vol. 6, no. 8, pp. 929 - 942, 2007.

[120] C. Ozturk, D.Karaboga and B. Gorkemli, "Probabilistic dynamic deployment of wireless sensor networks by artificial bee colony algorithm," *Sensors,* vol. 11, no. 6, pp. 6056 - 6065, 2011.

[121] M. Erdelj, T. Razafindralambo and D. S. Ryl, "Covering Points of Interest with Mobile Sensors," *IEEE Transaction on Parallel and Distributed Systems,* vol. 24, no. 1, pp. 32-43, 2013.

[122] A. P. Bhondekar, R. Vig, M. Singla, C. Ghanshyam and P. Kapur, "Genetic Algorithm Based Node Placement Methodology for Wireless Sensor Networks," in *Proceedings of the International Multi Conference of Engineers and Computer Scientists*, Hong Kong, China, 2009.

[123] Y. Zou and K. Chakrabarty, "Sensor deployment and target localization in distributed sensor networks," *ACM Trans. Embedded Comput. Syst,* vol. 3, no. 1, pp. 61-91, 2004.

[124] C. H. Wu, K. Chuan and Y. C. Chung, "A Delaunay Triangulation Based Method for Wireless Sensor Network Deployment," in *Proceedings of the 12th International Conference on Parallel and Distributed Systems (ICPADS'06)*, Minneapolis, MN, 2006.

[125] N. Kukunuru, B.R.Thlla and R. L. Davuluri, "Sensor deployment using particle swarm optimization," *International Journal of Engineering Science and Technology,* vol. 2, no. 10, pp. 5395-5401, 2010.

[126] X. Wang, J. Ma, S. Wang and D. Bi, "Distributed Particle Swarm Optimization and Simulated Annealing for Energy-efficient Coverage in Wireless Sensor Networks," *Sensors,* vol. 7, no. 5, pp. 628-648, 2007.

[127] G. Wang, L. Guo, H. Duan, L. Liu and H. Wang, "Dynamic Deployment of Wireless Sensor Networks by Biogeography Based Optimization Algorithm," *Journal of Sensor and Actuator Networks,* vol. 1, no. 3, pp. 86-96, 2012.

[128] O. Banimelhem, M. Mowafi and W. Aljoby, "Genetic Algorithm Based Node Deployment in Hybrid Wireless Sensor Networks," *Communications and Network,* vol. 5, no. 4, pp. 273 - 279, 2013.

[129] J. Chen and H. Qian, "Node Deployment Algorithm Based on Viscous Fluid Model for Wireless Sensor Networks," *The Scientific World Journal,* vol. 2014, no. Article ID 350789, 2014.

[130] I. Maleki, S. Khaze, M. Tabrizi and A. Bagherinia, "A new approach for area coverage problem in wireless sensor networks with hybrid particle swarm

optimization and differential evolution algorithms," *International journal of Mobile Network Communications & Telematics*, vol. 3, no. 6, pp. 61-75, 2013.

[131] S. S. Dhillon and K. Chakrabarty, "Sensor placement for effective coverage and surveillance in distributed sensor networks," *in Proc. of IEEE Wireless Commun. Netw. Conference,* vol. 3, pp. 1609-1614, 2003.

[132] F. Nematy and N. Rahmani, "Using voronoi diagram and genetic algorithm to deploy nodes in wireless sensor network," in *Proceeding of international conference on soft computing and software engineering(SCSE'13),* San Francisco,CA,USA, 2013.

[133] X. Yan, Y. Zhu and W. Zou, "A hybrid artificial bee colony algorithm for numerical function optimization," in *Proc. IEEE Int. Conf. on Hybrid Intelligent Systems*, Melacca , pp. 127-132, 2011.

[134] G. Wang, G. Cao and T. L. Porta, "Movement assisted sensor deployment," *IEEE Transactions on Mobile Computing,* vol. 5, no. 6, pp. 640-652, 2006.

[135] R. Özdag and A. Karci, "Sensor Node Deployment Based on Electromagnetism – Like Algorithm in Mobile Wireless Sensor Networks," *International Journal of Distributed Sensor Networks,* vol. 2015, no. Article ID 507967, 2015.

[136] W. H. Liao, Y. C. Kao and Y. S. Li, "A Sensor Deployment Approach Using Glowworm Swarm Optimization Algorithm in Wireless Sensor Networks," *Expert Systems with Applications,* vol. 38, pp. 12180 - 12188, 2011.

[137] C. Murthy, S. Ram and B. S. Manoj, "Ad Hoc Wireless Networks: Architectures and Protocols," 2004.

[138] L. A. Villas, A. Boukerche, H. S. Ramos, D. Oliveira, R. B. D. Araujo and A. A. F. Loureiro, "DRINA: A Lightweight and Reliable Routing Approach for In-Network aggregation in Wireless sensor networks," *IEEE Transactions on Computers,* vol. 62, no. 4, pp. 676-689, 2013.

[139] C. E. Perkins and P. Bhagwat, "Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers," *SIGCOMM,* vol. 24, no. 4, pp. 234 - 244, 1994.

[140] S. K. Sahani and K. Kumar, "Multi-routing in wireless sensor networks using an ant colony optimization (ACO)," *International Journal of Computer Networking, Wireless and Mobile Communications,* vol. 3, no. 3, pp. 87-98, 2013.

[141] H. Frey, S.Ruhrup and I.Stojmenovic, "Routing in wireless sensor networks," in *Guide to Wireless Sensor Networks,* Berlin,Germany, pp. 81 - 111, 2009.

[142] H. Keshavarz, R. Noor and E. Mostajeran, "Using routing table flag to improve performance of AODV routing protocol for VANETs environment," in *Proceeding of 9th International conference on Computing and Information Technology(IC2IT'13)*, pp. 73-82, 2013.

[143] C. E. Perkins and E. M. Royer, "Ad-hoc on-demand distance vector routing," in *Mobile Computing Systems and Applications, 1999. Proceedings. WMCSA '99. Second IEEE Workshop on*, New Orleans, LA, pp. 90-100, 1999.

[144] G. Caro and M.Dorigo, "AntNet: Distributed stigmergetic control for communications networks," *Journal of Artificial Intelligence Research,* vol. 9, pp. 317-365, 1998.

[145] G. D. Caro, F. Ducatelle and L. Gambardella, "AntHocNet: An ant-based hybrid routing algorithm for mobile ad hoc networks," in *Proc. 8th Int. Conf. PPSN, pp. 461-470*, 2004.

[146] J. Wang, E. Osagie, P.Thulasiraman and R. K.Thulasiram, "HOPNET: A hybrid ant colony optimization routing algorithm for mobile ad hoc network," *Ad Hoc Networks,* vol. 7, no. 4, pp. 690-705, 2009.

[147] R. Vidhyapriya and P. T. Vanathi, "Energy efficient adaptive multipath routing for wireless sensor networks," *IAENG International Journal of Computer Science,* vol. 34, no. 1, p. 56, 2007.

[148] M. M. Goswami, R. V.Dharaskar and V. M. Thakare, "Fuzzy ant colony based routing protocol for mobile ad hoc network," in *International Conference on Computer Engineering and Technology,CCET'09*, Singapore, pp. 438-444, 2009.

[149] S. Banerjee, A. Majumdar, H. N. Saha and R. Dey, "Modified Ant Colony Optimization (ACO) based routing protocol for MANET," in *International Conference and Workshop on Computing and Communication (IEMCON),* Vancouver, BC,pp. 1-7, 2015.

[150] M. N. Jambli, W. B. W. M. Shuhaimi, H. Lenando, J. Abdullah and S. M. Suhaili, "Enhancement of AODV routing protocol in MASNETs," in *9th International Conference on IT in Asia (CITA),* Kota Samarahan, pp. 1- 6, 2015.

[151] A. M. Ortiz, F. Royo, T. Olivares, J. C. Castillo, L. Orozco-Barbosa and P. J. Marron, "Fuzzy-logic based routing for dense wireless sensor networks," *Telecommunication Systems,* vol. 52, no. 4, pp. 2687 - 2697, 2013.

[152] X. Cao, 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," *Proc. 4th Int. Conf. Natural Comput.*, vol. 7, pp. 599-603, 2008.

[153] F. Ducatelle, G. A. D. Caro and L. M.Gambardella, "Principles and applications of swarm intelligence for adaptive routing in telecommunications networks," *Swarm Intelligence,* vol. 4, no. 3, pp. 173-198, 2010.

[154] S. Okdem and D. Karaboga, "Routing in wireless sensor networks using an ant colony optimization(ACO) router chip," *Sensors,* vol. 9, no. 2, pp. 909 - 921, 2009.

[155] B. Kadri, M. Feham and A. Mhammed, "Efficient and secured ant routing algorithm for wireless sensor networks," *International Journal of Network Security,* vol. 16, no. 2, pp. 149 - 156, 2014.

[156] M. Zafri and H. Khotanlou, "A routing algorithm based on an ant colony ,local search and fuzzy inference to improve energy consumption in wireless sensor networks," *International Journal of Electrical and Computer Engineering,* vol. 3, no. 5, pp. 640 - 650, 2013.

[157] A. Abbasi and M. Younis, "A survey on clustering algorithms for wireless sensor networks," *Computer Communications,* vol. 30, pp. 2826 - 2841, 2007.

[158] K. Sohrabi, J. Gao, V. Ailawadhi and G. Pottie, "Protocols for self-organization of a wireless sensor network," *IEEE Personal Communications,* vol. 7, no. 5, pp. 16 - 27, 2000.

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

[160] S. Thilagavathi and B. G. Geetha, "Energy aware swarm optimization with inter cluster search for wireless sensor network," *The Scientific World Journal,* vol. 25, pp. 1-8, 2015.

[161] B. Alkazemi and C. K. Mohan, "Multi-phase discrete particle swarm optimiziation," in *Proceeding of Fourth International Workshop on Frontiers in Evolutionary Algorithms (FEA 2002),* Kinsale, Ireland , 2002.

[162] J. C. Bansal, H. Sharma, K. Arya and A. Nagar, "Memetic search in artificial bee colony algorithm," *Soft Computing,* vol. 17, no. 10, pp. 1911 - 1928, 2013.

AUTHOR'S BIOGRAPHY

RAJESH KUMAR YADAV
Assistant Professor
Department of Computer Science & Engineering
Delhi Technological University, Delhi, India
Email:rkyadav@dtu.ac.in

Rajesh Kumar Yadav received his Bachler of Engineering (B.E) in Computer Sc. & Engineering from G. B. Pant Engineering College, Pauri (H.N.B. Garahwal University), Uttara Khand (UK) in year 2000 and Master of Technology (M.Tech.) in Computer Science from Uttar Pradesh Technical University, Lucknow, Uttar Pradesh (UP) in the year 2006. Presently he is working as an assistant professor in the department of computer science and engineering, Delhi Technological University, Delhi, India-110042.His area of research includes wireless sensor networks and Mobile ad-hoc networks.