ASSESSMENT OF DYNAMIC CLUSTERING IN WIRELESS SENSOR NETWORKS FOR THE OPTIMIZATION OF ENERGY CONSUMPTION

by

Aslı Balta

B.S., Electronics and Communications Engineering, Yıldız Technical University, 2008

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Electrical and Electronics Engineering
Boğaziçi University
2011

ACKNOWLEDGEMENTS

I would like to show my deepest gratitude to my advisor Prof. Emin Anarım. He has been an excellent supervisor, in terms of professional support, encouragement and countless helps for all difficulties.

This thesis would not have been possible without the help and support of Tolga Kurt. Since the beginning of the study he has consistently provided invaluable help and advice.

My parents has always been extremely supportive in many ways all my life and they deserve a special mention. l also would like to thank to my friend Halim Terzioglu for his patience and encouragement during all the period of this thesis work.

I also thank TÜBİTAK for supporting this thesis with BİDEB 2210.

ABSTRACT

ASSESSMENT OF DYNAMIC CLUSTERING IN WIRELESS SENSOR NETWORKS FOR THE OPTIMIZATION OF ENERGY CONSUMPTION

Wireless sensor networks are the networks that contain many sensors which have different type of abilities such as sensing, communicating or processing. In recent years there are many applications based on these kind of networks. Since most of the applications in wireless sensor networks require energy consumption, the energy level of any single node becomes important. There are different methods in literature which provides efficient energy consumption for wireless sensor networks. Wireless network clustering is one of the major methods which has been used in many applications for energy optimization. As there are many nodes in the networks it is aimed to provide an adaptive energy consumption with clustering methods. LEACH is one of the most popular algorithm for clustering which is based on random clustering method. In this algorithm there are two different type of nodes as cluster heads and cluster members. Communication is provided via these cluster heads. Therefore LEACH algorithm reaches its main purpose by providing dynamic clustering and adaptive energy consumption.

The main clustering algorithms are analyzed and drawbacks of the LEACH algorithm are determined in this thesis. New clustering algorithms are provided in order to prevent undesired cases of LEACH clustering algorithm. It is shown in this thesis that random methods are beneficial in clustering but not adequate enough to reach the desired results. Improving the random algorithms with different constraints and methods reaches better results in case of energy consumption.

ÖZET

KABLOSUZ ALGILAYICI AĞLARDA ENERJİ TÜKETİMİNİN ENİYİLENMESİ İÇİN DİNAMİK KÜMELEMENİN İNCELENMESİ

Kablosuz algılayıcı ağlar sezme, haberleşme veya işleme gibi değişik özelliklere sahip algılayıcıları içeren ağlardır. Son yıllarda bu tip ağlar üzerine pek çok uygulama görülmektedir. Bu uygulamaların pek çoğu enerji tüketimi gerektirdiği için herbir düğümün enerji seviyesi önemli hale gelmektedir. Literatürde verimli enerji tüketimi sağlayan değişik metotlar vardır. Enerji optimizasyonu için kullanılan en önemli metotlardan biri kablosuz algılayıcı ağları kümelere ayırmaktır. Bu ağlarda pek çok düğüm olduğu için kümelemeyle adaptif enerji tüketimi sağlanmaktadır. LEACH rastgele kümelemeye dayanan en popüler algoritmalardan biridir. LEACH algoritmasında küme başkanı ve küme elemanı olmak üzere iki ayrı düğüm vardır ve haberleşme bu küme başkanları üzerinden sağlanmaktadır. Bu sayede de LEACH algoritması dinamik kümeleme kullanarak ve bütün düğümlere eşit şans vererek esas amacına ulaşmaktadır.

Bu tezde temel kümeleme algoritmaları incelenmiş ve LEACH algoritmasının olumsuz tarafları belirlenmiştir. LEACH algoritmasının yol açtığı olumsuz durumların önüne geçmek amacıyla yeni kümeleme algoritmaları sunulmuştur. Bu tezde rastgele seçim yöntemlerinin kümeleme açısından faydalı olduğu fakat istenilen sonuçlara ulaşmak için yeterli olmadığı ve rastgele seçim yöntemlerinin değişik sınırlamalarla geliştirilebileceği gösterilmiştir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

C_{prob}	Small constant that is defined by algorithm
$C_{h_{prob}}$	Probability of a single node to become a cluster head
d	Distance between nodes
E_{elec}	Radio dissipation factor
E_{ncur}	Current energy of a node
E_{nmax}	Initial energy of a node
E_{res}	Residual energy level of a single node
E_{rx}	Receiver's energy
E_{tx}	Transmitter's energy
po	Centroid of the cluster
p	Desired percentage of cluster heads
P_t	Transmission power
r	Number of round
r_s	Number of rounds for a node that had not become a cluster head
tc	Short iteration in HEED algorithm
$T_{n_{new}}$	New threshold value
T_n	Threshold value for cluster head selection

1. INTRODUCTION

1.1. Wireless Sensor Networks

Many systems of scientific interest can be represented as networks, sets of nodes or vertices joined in pairs by lines or edges such as internet, world wide web, food webs, neural networks, metabolic networks. The study of networked systems has a history stretching back several centuries, but it has experienced a particular surge of interest in the last decade, especially in the mathematical sciences [1] and recently wireless sensor networks also have become an important area in research and science.

Wireless sensor networks are the networks containing intelligent sensors which have different properties such as sensing, processing or communicating. Due to this reason sensors are usually equipped with data processing and communication capabilities. The sensing circuit is used to measure the parameters in the environment and transform these measures to electrical signal. Each sensor has an onboard radio that can be used to send the collected data to interested parties [2], [3]. There are different applications based on wireless sensor networks related with the abilities of these sensors. These wireless sensor networks are mainly used for monitoring, data gathering and communicating. Due to the difference between these applications sensors in different networks may require different properties.

In case of monitoring, the sensor nodes usually need to collect and send the data of some predefined parameters. Since they need to communicate with each other or an administrator these sensors need the ability of communication and data gathering. For example in a habitat monitoring application, user needs to monitor some parameters such as temperature and humidity. Since this network probably will contain large number of sensors data gathering will become a bottleneck for the application. There are different approaches for the problem of data gathering mainly based on data aggregation.

A network of sensors can also be employed to gather meteorological variables such as pressure.These measurements can be then used in preparing forecasts or detecting harsh natural phenomena. In disaster management situations such as earthquakes, sensor networks can be used to selectively map the affected regions directing emergency response units to survivors. In military situations sensor networks can be used in surveillance missions and can be used to detect moving targets, chemical gases, or the presence of micro-agents [4].

In order to design good protocols for wireless microsensor networks, it is important to understand the parameters that are relevant to the sensor applications. Sensor networks may contain hundreds or thousands of nodes, and they may need to be deployed in remote or dangerous environments. In order for the efficiency of application deployment should be easier. These networks should funtion as long as they can and the sensors are usually not able to recharge batteries. Therefore hardware and software structures should be deployed with energy consideration [5].

The wireless sensors in the networks can have different properties accrording to the applications as given above. Most of the applications based on sensor networks require communication between nodes and in order to provide communication these sensors usually have a radio transmitter. Moreover communication between any nodes require energy consumption, therefore the energy of a single node in the network becomes so limited.

Recently energy consumption in wireless sensor networks has become a hot research area, there are different algorithms based on different approaches for the optimization of energy consumption.

In addition to the problem of energy consumption one major problem of wireless sensor networks is the difficulty of monitoring these networks. There are different researchs in literature for the monitoring of sensor networks. The structure of these protocols differs with their different aims; discover failed nodes [6],[7], compute the coverage [8],[9], determine the remaining energy level [10] or topological mapping of the network [11]. The authors in [12] provides a monitoring tool which continuosly aggregates and computes different properties of the networks such as loss rates, energy levels or packet loss.

In recent years, ad hoc wireless networks have attracted a lot of attention due to multiple potential applications, both civilian and military. Various approaches have appeared in the literature for the study of this type of networks [13]. Some of these studies are based on clustering schemes for wireless networks and recently there has been many applications on this area.

1.2. Wireless Network Clustering

In the continuing furry of research activity within physics and mathematics on the properties of networks, a particular recent focus has been the analysis of communities within networks.

Detection of these communities has become a new interest in research areas and there are different methods based on different theories such as graph theory. Many of these scientific studies has shown that community detection in wireless networks gives rise to different applications. Wireless network clustering is one of the most used technique in wireless sensor networks in order to prolong network lifetime and increase the efficiency of data gathering. Since these networks usually have large scale of nodes the communication based on clustering namely hiearchical routing results an effective communication scheme.

In a clustered network there are usually two kinds of nodes such as clusterheads and cluster members. Once the clusters and clusterheads are determined hierarchical routing can take place. In this routing scheme cluster members only need to communicate with their clusterheads, in some applications they also may need to communicate with the members within same cluster. On the other hand cluster heads need to be in communication both with other clusterheads and the members of their own cluster. In case of any communication, cluster members just need to send their data to the cluster head, they do not need to know whole topology and this provides data aggregation and decrease in energy consumption. Cluster heads will need to do all other things left such as finding destination address, computing the shortest path and sending the message via the shortest path. In order to design a more efficient network and prolong the networks life time the selection of cluster heads becomes very important.

Clustering has so many advantages in wireless sensor networks such as network scability. It can also localize the route set up within the cluster and thus reduce the size of the routing table stored at the individual node [14]. Moreover clustering can conserve communication bandwidth since it limits the scope of inter-cluster interactions to CHs and avoids redundant exchange of messages among sensor nodes [15]. Furthermore, a CH can aggregate the data collected by the sensors in its cluster and thus decrease the number of relayed packets [16]. Clustering also enables bandwidth reuse and can, thus, increase system capacity. Using clustering enables better resource allocation and helps improve power control [17].

Yang Qin and Jun He re-phrase [18] in 2004 about the hierarchical routing in ad hoc wireless networks. The purpose of this paper is to investigate the network throughput in ad hoc wireless networks under a cluster-based hierarchical routing scheme. In mobile ad hoc networks based on hierarchical routing protocols , all of the nodes do not need to know whole topology. Only some nodes called cluster-head nodes know the topology information and all the other nodes can simple send their packets to these cluster head nodes. The cluster head nodes will finish all the remaining things. The traffics will be concentrated in cluster-head nodes and thus the cluster-head node may become the bottleneck of an ad hoc network. They observe that the throughput of network with such hierarchical routing scheme will be smaller than the one in flat routing scheme [18].

The authors in [19] have proposed a different kind of clustering structure which aims to organize whole network into smaller clusters an subclusters. In case of this clustering structure sensor nodes deployed in the wide area, will form many cluster groups for efficient network organization, where each cluster group contains sensor nodes in majority, one cluster head and one node leader. The main role of node leader is to gather and aggregate the sensor data from other sensor nodes in the same cluster group. The cluster head will then forward the aggregated data coming from the node leader, to the base station either directly or through other cluster heads. The authors also provide a fault tolerant clustering algorithm with this structure. In case of any fault in sensors, sensor data will be unreachable to calculate the critical threshold value, for predicting the event in the related application. Even if the sensor node fails, it is possible to give an approximately predicted sensor data from the same region, in the same sensor reading time interval. Thus the sensor information from the failed sensor node can be approximated from the geographically nearest sensor [19].

Recently there are many different approaches for clustering and cluster head selection. Different algorithms give rise to different improvements. Since wireless sensor networks become more and more important everyday clustering techniques seem to be improved in next years.

1.3. Wireless Sensor Network Clustering Applications

In recent years, ad hoc wireless networks have attracted a lot of attention due to multiple potential applications, both civilian and military. Various approaches have appeared in the literature for the study of this type of networks[13]. Some of these studies are based on clustering schemes for wireless networks and recently there has been many applications on this area.

In 2001, Beongku and Papavassiliou studied on an architecture for supporting geomulticast services in mobile ad hoc wireless networks. Since in ad hoc networks there is no infrastructure as in cellular networks they use mobilitiy based clustering to cluster the various mobile nodes [20]. For the implementation first wireless network clustering is made and a head cluster is chosen for each group, their network structure consists of elements similar to cellular networks. For example cluster heads act as base station in a cellular network and mobile nodes are similar to mobile station. Using this structure as basis they propose a direction guided routing protocol which creates a cluster head based limited mesh structure within a guided region.

Another paper based on wireless network clustering is published by Taghi,et.al. in 2005 about the intrusion detection in wireless networks using clustering techniques with expert analysis. The authors assume that since intrusive activity is relatively very small compared to normal activity, normal traffic records are generally very dominant in the collected data. Consequently, the largest cluster will largely be composed of normal instances given the hypothesis that many of the normal instances will have similar attributes. It is then very likely that intrusive instances would belong to clusters that are farthest to the largest cluster. Therefore the steps of the algorithm is as follows; Finding the cluster with the most number of instances, i.e., largest cluster, and label it normal. Let the centroid of this cluster be po. Sorting the remaining clusters in ascending order of the distance from each cluster centroid to po. Selecting all clusters that have a distance (to po) greater than qD, and label them as intrusive, where q is selected by an expert upon inspection of empirical data. Labeling all the other instances as normal [21].

In 2008, Z. Wang et.al. has published a paper about a position based clustering technique for ad hoc intervehicle communication [22]. In an intervehicle system two passing vehicles can exchange data or vehicles can act as routers and transmit the data to another vehicle. With this principle, highly efficient accident warning systems are possible; cars involved in an accident can send warning messages back over a predefined number of other vehicles, thus avoiding motorway pileups and enhancing the traffic safety. The importance of intervehicle communication for emergency is also described in [23] as in case of emergency situations, it is paramount to be able to forward important information as soon as possible. Such emergency information could be originated by first responders or locally by cars on highway and could be directed to the whole network (broadcast) to part of the network (broadcast to specific highways) or to special centers such as police, emergency rooms, hospitals, etc.

Traditional topology management approaches are not designed for the earlier vehicular environment so the authors in [22] propose a position-based prioritized clustering (PPC), a new topology control method for vehicular ad hoc networks This proposal incorporates position information into a novel hierarchical clustering technique. The new network topology enables nodes to move during cluster setup and maintenance. The dismiss threshold controls the cluster reconfiguration. This paper performs basic mathematical analysis of PPCs performance under some assumptions. Simulations show that this new technique has flexibility and stability.

Monitoring of wireless sensor networks is one of the major problems that is given in the previous section. The authors in [24] provides a hierarchical monitoring application based on the clustering. The application is mainly provided to collect the residual energy information of the sensors continuosly to construct an energy map at the base station. However due to the large number of sensors this data collection requires high energy consumption and in order to provide an energy efficient algorithm clustering techniques are used in this work. At the begining whole network is seperated into clusters and for each cluster some nodes called cluster heads are determined to act as aggregator in data collection process. At the same time, a topology tree is constructed which consists of these head nodes and some bridging delivery nodes between two adjacent clusters. Based on the energy information collected, a set of polygons which represent the contours of different energy levels are produced independently for each cluster. The topology tree is then used to collect the energy graphs from leaf nodes to the base station. Therefore the cost for message transmission is reduced by using clustering and in network aggregation [24].

1.4. Outline of the Thesis

This thesis addresses the problem of energy consumption in wireless sensor networks with network clustering techniques. Rest of the thesis is organized as follows; Some network clustering techniques are briefly explained in Chapter 2. Chapter 3 includes the details of hierarchical routing type. The details of the new clustering algorithm and simulation results are given in the Chapter 4. Comparison has been done between new clustering algorithms in the same chapter and final decision for clustering algorithm is also given in Chapter 4. In this chapter a deep analyze has been done on new algorithms and the applicability of algorithms in real networks. And finally Chapter 5 concludes the thesis.

2. WIRELESS NETWORK CLUSTERING TECHNIQUES

Wireless network clustering has become a popular reseach area within the increase of applications based on wireless sensor networks. In literature there are different approaches to this problem. The main clustering techniques can be listed as LEACH(Low Energy Adaptive Clustering Hierarchy) , HEED(Hybrid Energy Efficient Distributed) etc.

2.1. LEACH

LEACH [25] is the first clustering protocol proposed for periodical data gathering applications in WSNs. In LEACH, the nodes organize themselves into local clusters, with one node acting as the cluster head. All non-cluster head nodes transmit their data to the cluster head, while the cluster head node receives data from all the cluster members, performs signal processing functions on the data,and transmits data to the remote BS [5].

LEACH provides an energy adaptive clustering algorithm by a dynamic topology with cluster heads. The operation of LEACH is divided into rounds. Each of these rounds consists of a set-up and a steady-state phase.During the set-up phase cluster-heads are determined and the clusters are organized. During the steady-state phase data transfers to the base station occur. With this algorithm cluster heads and clusters change periodically within a time called round. In each round a threshold is determined and with this threshold degree every node decides if they are cluster head or not for that round. In order to do this every node determines a random number for each round and if this random number is smaller than that rounds threshold then this node becomes a cluster head for that round. In LEACH algorithm p defines the desired percentage of cluster heads in a single round and due to the aim of enery optimization if a node becomes a cluster head in first round then it will not be selected as a cluster head in the next 1/p rounds. Therefore it is certain that all of the nodes can become cluster head but only once in each 1/p round and only nodes that have not already been cluster head recently will have a chance to become a cluster head. The algorithm of cluster head selection is based on the below formula;

$$T_n = \frac{p}{1 - p\,(r \bmod 1/p)} \tag{2.1}$$

In the above Formula 2.1 T_n defines the threshold value to become a cluster head, p is the desired percentage of clusterheads in the network and r stands for the number of round. After cluster head selection every node determines its cluster according to the closest cluster head.

Once the nodes have elected themselves to be cluster heads using the algorithm, the cluster head nodes must let all the other nodes in the network know that they have chosen this role for the current round. To do this, each cluster head node broadcasts an advertisement message. This message is a small message containing the nodes ID and a header that distinguishes this message as an announcement message. Each non-cluster head node determines its cluster for this round by choosing the cluster head that requires the minimum communication energy, based on the received signal strength of the advertisement from each cluster head. Assuming symmetric propagation channels for pure signal strength, the cluster head advertisement heard with the largest signal strength is the cluster head that requires the minimum amount of transmit energy to communicate with. After each node has decided to which cluster it belongs, it must inform the cluster head node that it will be a member of the cluster. Each node transmits a join-request message back to the chosen cluster head.This message is again a short message, consisting of the nodes ID and the cluster heads ID [25].

In case of any communication between two nodes, the transmitter node just needs to send the message to its cluster head and the remaning parts will be completed by the cluster head. Once the cluster head receives the message it will send the message to receivers cluster head and this node will send the message to the final receiver node.

Therefore an efffective way of energy consumption is provided and moreover since all of the nodes does not need to know whole topology this communication structure will decrease the complexity.

On the other hand in case of direct communication every node needs to send their messages directly to the BS or another receiver. If the receiver is far away from the node then transmission will require higher amount of energy therefore the transmitter and receiver nodes both will lose a large amount of their energy. This will cause nodes to quickly drain their batteries and reduce the networks lifetime. However this type of communication may also be acceptable if the nodes are located close to each other or BS.

According to the authors in [25] LEACH can achieve over a factor of 7 reduction in energy dissipation compared to direct communication. In addition to reducing energy dissipation, LEACH successfully distributes energy-usage among the nodes in the network such that the nodes die randomly and at essentially the same rate [25]. Although LEACH provides an adaptive energy consumption and it increases the efficiency of wireless sensor networks there are still some problems with energy consumption and data aggregation.

Since every node can become a clusterhead in LEACH algorithm it sometimes may result with undesired topologies. In some cases border nodes can become clusterhead and in this case the higher distance between cluster heads and cluster heads members increases energy consumption and results an unefficient network. On the other hand distance between the cluster heads also becomes important for efficiency, one of the undesired state in clustering topology is the small distance between cluster heads. In order for efficiency to be higher, distance between the cluster heads should be adequate enough for occurance of two different clusters.

As a result LEACH provides efficiency for wireless sensor networks however it sometimes may cause inefficient topologies. Some clustering results for good case and bad case are given below in Figure 2.1 and Figure 2.2. In these figures clusters are given with different markers and cluster heads are also marked with square. As it can be seen from Figure 2.1 there are four different clusters and cluster heads are not border nodes. Additionaly the distance between the cluster heads are adequate according to the total size of network. Figure 2.2 shows the bad case scenario for LEACH. In this case there are six clusters but since two cluster heads are located so close to each other, the selection of these cluster heads may cause inefficiency. Recently there are different research in literature providing different solutions and improvements for LEACH algorithm.

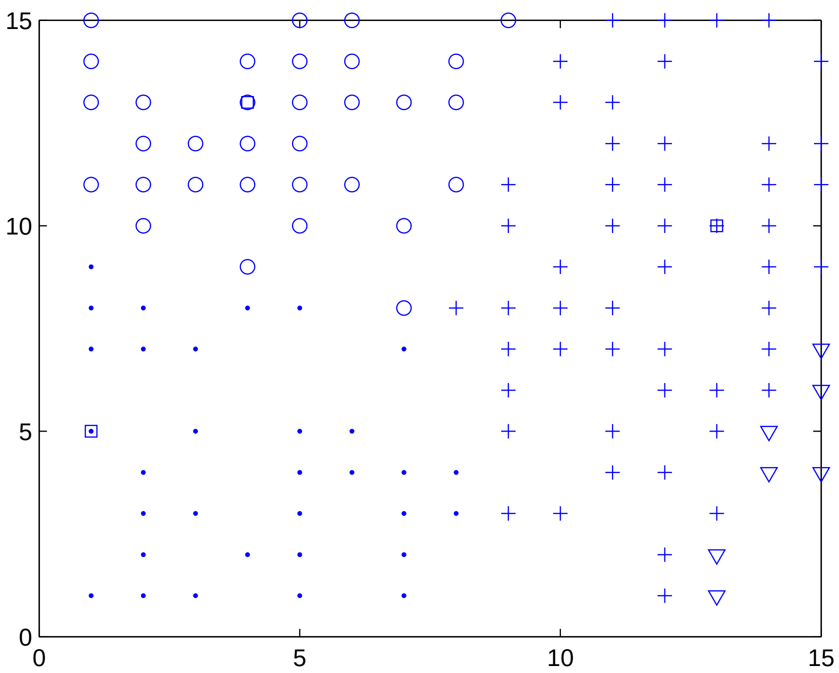

Figure 2.1. Good case of LEACH.

M. J. Handy, et all has published a paper about an improved version of LEACH algorithm. They point out some disadvantages of LEACH such as the probability of the nodes with lower level energy becoming cluster heads. Their first approach for increasing the lifetime of the network is taking into account the remaining energy of nodes in cluster head selection. In case of this calculation a node has to have enough energy to become a cluster head therefore low energy nodes will not have the right to become cluster head. In order to apply this approach to the algorithm they modified

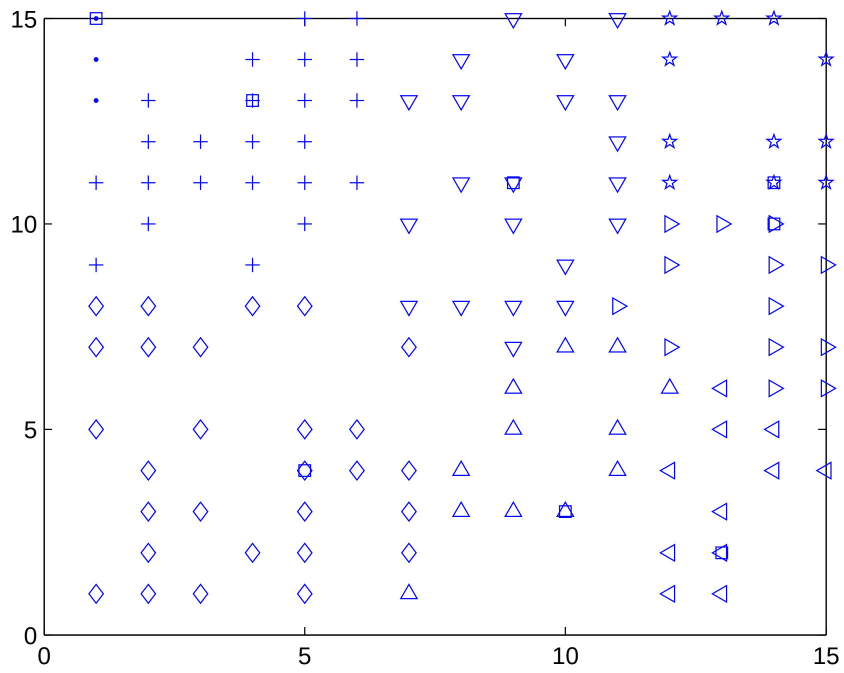

Figure 2.2. Bad case of LEACH.

LEACH formulation as below;

$$T_n = \frac{p}{1 - p\,(r \bmod 1/p)} \frac{E_{ncur}}{E_{nmax}} \tag{2.2}$$

E_{ncur} is the current energy of the node and E_{nmax} is the inital energy of that node. Therefore they represent the energy level with the coefficient Encur/Enmax. This approach leads higher energy level nodes to become cluster heads and simulation results show that improvement in efficiency can be provided. However there are still some disadvantages also in this case, after certain number of rounds network becomes stuck. Since after some certain rounds most of the nodes will have low level of energy, the threshold for becoming a cluster head will become too low. Although there will be still some nodes which have enough energy to send data due to the low energy level of threshold the network will already become stuck.

The authors also provides a solution for this problem in [26] by a further modification of the formulation of cluster head threshold equation. The equation is modified as in 2.3 with a coefficient for the nodes that has not become a cluster head in 1/p rounds.

$$T_{n_{new}} = T_{n_{LEACH}} \left(\frac{E_{ncur}}{E_{nmax}} + r_s p - \frac{r_s p E_{ncur}}{E_{nmax}} \right) \tag{2.3}$$

In equation 2.3 r_s is the number of rounds for a node that had not become a cluster head, if this number reaches 1/p then the formula will be modified the older version as in LEACH then $T_{n_{new}}$=$T_{n_{LEACH}}$. Therefore remaining nodes will have chance to become clusterhead and in other cases rs will be set to 0 in order to achieve modified formula.

By this modification authors has solved the problem of stuck network and also they have reached a more effective energy consumption than LEACH. With these modifications a 30 percent of increase in lifetime of microsensor networks can be accomplished.

The authors in [26] has discussed two modifications of LEACHs cluster-head selection algorithm. Furthermore, an important quality of a LEACH network is sustained despite the modifications: For the deterministic selection of cluster-heads only local and no global information is necessary. The nodes themselves determine whether they become cluster-heads. A communication with the base station or an arbiter-node is not necessary.

Another variation of LEACH algorithm is given in [27] with HCR protocol. The main objective of the HCR protocol is to generate energy-efficient clusters for randomly deployed sensor nodes, where each cluster is managed by a set of associates called a head-set. Using round-robin technique, each associate member acts as a cluster head.

Cluster formation part of HCR is very similar to LEACH except the selection of head set nodes. Once the clusters and cluster heads are determined cluster heads select a predefined number of head set nodes according to the signal strength of acknowledgement messages. At the end of election phase, each head-set member checks if it has sufficient energy for next round. If the energy of any head-set member falls below the given threshold value, it is removed from the head-set; the remaining head-set members update their schedules accordingly. At the end of round, all the clusters are not destroyed; however, each cluster is retained for the number of rounds equal to the head-set size [27]. Therefore nodes of the clusters with the head set size 1 can become candidate for the next round however the nodes of cluster whose head set size is greateer than 1 do not participate in the next election. This clustering structure reduces the number of elections and provides more efficient clustering.

2.2. HEED

The authors in [28] has approached a new clustering algorithm based on some probabilistic equations. In this algorithm it is assumed that nodes have no specialities such as having a GPS. The main approach is to cluster all the nodes in an equal way which is based on probability. As the other clustering techniques HEED algorithm also aims to prolong the network lifetime and increase the efficiency. In order to compare network time they defined a certain value as the first or last node depletes its energy. The main factor in the probabilistic approach is the residual energy of nodes.

In HEED algorithm every node is exactly mapped to one cluster and this node has to be able to communicate with the clusterhead via single hop. The tranmission ranges and energy levels are classified and defined as inter-cluster transmission range, inter-cluster power, intra-cluster transmission range and intra cluster power. Inter cluster transmission range is higher and inter transmission requires more energy than intra cluster transmission as expected.

Cluster head selection is mainly based on two different approaches that are about energy level and cost. In order to consider the energy levels of nodes for cluster head selection authors define an initial set including high energy level nodes. Therefore it is prevented for low energy nodes to become cluster head. The second parameter cost is used to break ties between nodes. If two different nodes in the same intra cluster transmission range sends their willingness to become a cluster head a tie occurs between these two nodes.

The HEED algorithm is mainly based on probability of being a cluster head which is given with the following equation 2.4, all the nodes set their initial probability to become a cluster head as Chprob;

$$C_{h_{prob}} = \frac{C_{prob} E_{res}}{E_{max}} \tag{2.4}$$

Chprob is the probability of a single node to become a cluster head, Cprob is the small constant that is defined by algorithm. E_{res} stands for the residual energy level of the single node and Emax is the reference maximum energy. Cprob is used to limit the number of initial cluster head announcements. During each iteration, a node arbitrates among the cluster head announcements it has received to select the lowest cost cluster head. If it has not received any announcements, it elects itself as a cluster head with probability Chprob. If successful, it sends an announcement indicating its willingness to become cluster head. The node then doubles its probability Chprob, waits for a short iteration interval tc, and begins the next iteration. A node stops this process one iteration after its Chprob reaches 1. If a node elects to become a cluster head, it raises its transmission power to Pt for inter-cluster communication [28].

HEED provides an efficient clustering algorithm based on probabilistic to increase the network lifetime. There are some different approaches on HEED to increase efficieny. O. Younis et all has provided an improved algorithm IHEED which is mainly based on HEED. This algorithm integrates node clustering and multi hop routing in order to increase efficiency of network.

One of the most important challange in IHEED is integration of clusters in data aggregation trees without degrading path quality [29]. In this topology only cluster heads are used to construct the aggregation tree, since cluster heads will be distributed well even if the nodes are not well distributed path quality will be higher inter cluster level. Figure 2.3 shows the general structure of aggregation tree with IHEED clustering. Another challange that IHEED provides is the estimation of remaining energy

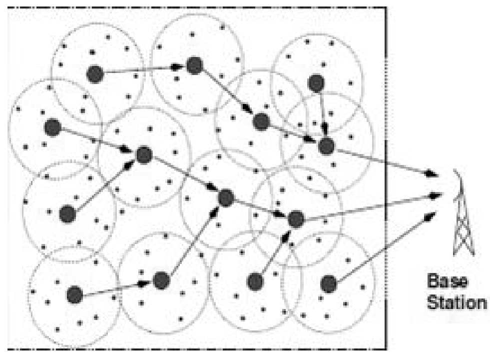

Figure 2.3. Spanning tree of cluster heads in IHEED [26].

of a node. Although this estimated energy level is not always guranteed it results more efficient clusters. The method is used to estimate energy level is credit point system which is already used in previous applications. It is a simple approach that compiles known physics formula, and converts Joule consumption into point deduction for simpler computations.

Simulation results show that IHEED reaches its goal of prolonging the network life time by integrating clustering to data aggregation. The simulations on a sensor network testbed demonstrate that by using clustering and data aggregation, the network lifetime is prolonged by a factor of 2 to 4, the number of successful transmissions is almost doubled, and the maximum overhead is reduced to less than half [28].

3. HIERARCHICAL ROUTING

The main concern of networks is to provide communication between two different nodes as one of them is originator of message and the other one is receiver. In case of existence a direct link between originator and router this routing will be simple however the networks in real life are usually too large to have direct link between all nodes. Therefore communication between two nodes becomes more complex.

Communication within a network can be thought as a traffic in a city. Traffic jam has already become a major problem for big cities and there are different solutions and new traffic models to solve this problem. If all the cars try to use the shortest way to reach the destination it will cause all of them to wait in traffic jam especially in rush hours. According to some traffic models there may be some alternative ways which do not provide the shortest path but may provide shortest time to reach destination. If all of the cars use different alternative ways the congestion in the traffic will be over. Today many navigator systems are mainly based on this idea. In case of networks it can be analyzed as a routing algorithm which provides the path with lowest cost.

The authors in [30] has provided a a dynamic vehicle routing system which is based on Ant Based Control algorithm. This routing system guides indivual drivers through city. Simulations has shown that this algorithm is very suitable for decentralized dynamic route optimization where dynamic data changes frequently. However it is shown that this algorithm cannot provide efficient results for big networks which include several cities and motorways between these cities [31]. In this case it is necessary to detect some communities, map them and provide routing based on these communities.

Hierarchical routing provides better solution for the case of larger networks that include several cities and connection between these cities. The main idea is to split up to network to small groups. In case of this traffic model every city can be mapped as a small network and the connection between cities belongs to the other network.

The resulting hierarchical network consists of two levels, the abstract level and the detailed level. The lower detailed level is the abstract level. On the abstract level there is only one network, the motorway network. To it belong all motorways and motorway intersections is given in Figure 3.1 . The detailed level consists of several independent city networks, one for each sector that contains a city network. To a city network belong all streets and intersections of the sector and also the surrounding motorways with their intersections is given in Figure 3.2. By splitting up the network in this way a hierarchical traffic network based on the human's behaviour to plan a route is created [31].

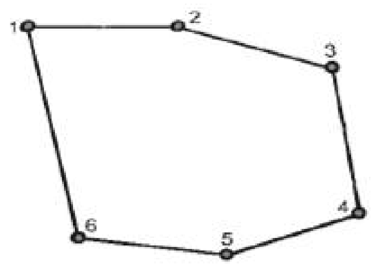

Figure 3.1. Global network [31].

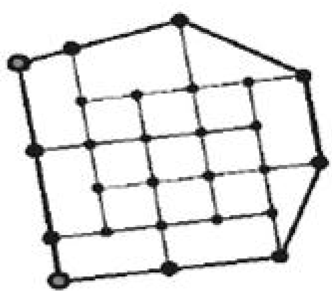

Figure 3.2. A sector network [31].

4. SIMULATION RESULTS FOR THE CLUSTERING

There are different types of network clustering algorithms in the literature and these algorithms usually differ with the aim of the algorithm. Many of the algorithms are improved depending on the requirements of the related application. The aim of this thesis is to provide an efficient clustering algorithm for energy optimization in wireless sensor networks. Since our aim is to optimize the energy consumption the best clustering algorithm that fits our aim is LEACH algorithm. However looking at a single round of LEACH it is obvious that a random cluster head selection does not always lead to robust network structures. The aim of this thesis is placed on the analyze and determination of drawbacks of LEACH and provide solutions to overcome the difficulties and prevent occurance of undesired cases in LEACH. The details of the improvement on LEACH will be given in this chapter.

LEACH provides a random clustering method but there are some constraints that effect the clustering algorithm such as definition of number of clusters in each round. In LEACH based algorithms a p value defined as desired percentage of cluster heads is used to calculate threshold values, therefore it directly effects the cluster head selection. As a first step the optimum value should be determined and the parameters that depends on p value should be selected to reach best results.

The authors in [26] showed that number of cluster heads has a big role in the efficiency of the optimization. They showed that 0 clusters or 100 percent clusters will give the same results with direct communication. Therefore there occurs a need to determine the optimum value for the number of clusters. This value mainly depends on the distribution of the nodes in a given network. Two networks with same number of nodes but different distributions may lead different results with same p value.

4.1. First Order Radio Model

In order to simulate energy efficiency of network it is assumed a simple model which is used in number of previous studies [5],[32],[33]. It is assumed in this model that radio dissipates Eelec=50nJ/bit and transmitter amplifier is eamp=100pJ/bit/m2. The following model is used to simulate the communication scenario;

$$E_{Tx}(k; d) = E_{Tx} - e_{elec}(k) + E_{Tx} - amp(k; d) \tag{4.1}$$

$$E_{Tx}(k; d) = E_{elec}k + e_{amp}kd^2 \tag{4.2}$$

$$E_{Rx}(k) = E_{Rx} - E_{elec}k \tag{4.3}$$

$$E_{Rx}(k) = E_{elec}k \tag{4.4}$$

For the communication scenario it is assumed that every node can send message to another one, and initially each node has an equal level of energy that is 0.5J.

For these parameter values, receiving a message is not a low cost operation; the protocols should thus try to minimize not only the transmit distances but also the number of transmit and receive operations for each message. The assumption is that the radio channel is symmetric such that the energy required to transmit a message from node A to node B is the same as the energy required to transmit a message from node B to node A.

4.2. Network Topology

Wireless sensor networks are usually located in larger geographical areas and can have different distributions. There are different papers in literature about the effect of network distribution to networks lifetime. he best results will be achived by locating the sensors with a controlable manner however this method does not seem possible in real networks. The most of the wireless sensor networks are randomly distributed and due to the large geographical areas of the networks it is not possible to locate all the sensors by hand.

In this thesis it is assumed that 200 nodes are randomly distributed on a geographical area of 4km2, the area is divided into squares of 0.01km2. Figure 4.1 shows the whole network topology for this random network which is based on uniform distribution.

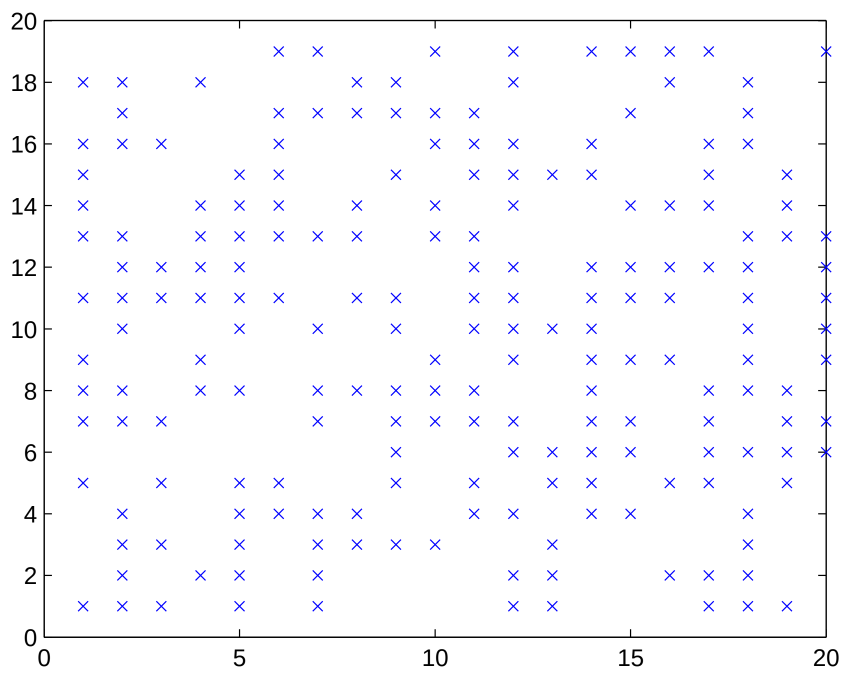

Figure 4.1. Network topology based on Uniform distribution.

4.3. Desired Percentage of Cluster Heads

LEACH provides an adaptive random clustering technique however there are still some constraints in the selection of cluster heads. The main parameter which directly effects the cluster head selection is desired percentage of cluster heads. Looking at the Formula 2.1 it is obvious that that any change in p value may effect the threshold value and so directly effects cluster head selection.

Desired percentage of cluster heads should be selected as optimum to provide best clustering results. This value mainly depends on the distribution of the nodes in a network. Actually if two networks have same number of nodes but distributions are different, same p values may result to dissimilar results. For example if the nodes in one network are close to each other it may be desired to have more number of clusters but in another network if nodes are distributed as not very close to each other and there are big distances between some groups of networks it may be desired less number of clusters.

In the second case distance between the cluster heads will be larger but distance within cluster will not be larger. Therefore in order to provide an effective clustering first the network distribution should be analyzed.

In order to select the best p value some simulations have been done on random network. For comparison of different p values, different measurement parameters are used such as number of dead nodes after a certain round. In this measurement dead node stands for the nodes that have no enough energy for any communication.

Simulation results show that usage of different p values has given different results that can be seen inTable 4.1. The analysis of Table 4.1 shows that increase in p number does not always leads better results and the optimum p value for this network seems to be 0.05. Moreover increase in the number of rounds also have an effect in p calculation. In first 1800 round it is seen that p values 0.1 and 0.05 give the same results but when the algorithm is applied for 2000 rounds it can again be analyzed that 0.05 percentage of cluster heads in this network will reach the optimum results.

Table 4.1. Different p values comparison.

P	Round	Number of dead nodes
0.04	1800	24
0.05	1800	22
0.1	1800	22
0.2	1800	25
0.04	2000	29
0.05	2000	26
0.2	2000	29

It can be obviusly seen from simulation results and Figure 4.2 that some larger p values will decrease the efficiency of network as well as small ones. Small p values will cause less clusters and therefore distance between clusters and within clusters will be larger than desired values. Moreover larger p values will result with more clusters in the network and in this case communication will be more complex and it may cause so many nodes to lose energy even for any message transmission. As it can be seen from the Figure 4.2 there is an optimum value for the number of cluster heads if there are less than optimum number of clusters some nodes will have to transmit their data very far, on the other hand if there are more than optimum number of clusters there will be more cluster heads in each round and it will cause more nodes to use energy for each round depending on the communication structure.

The best results of clustering is expected to get when the nodes in a network are located by hand, however the wireless sensor networks in real life are not always suitable for this type of networking. If the position of the nodes are previously determined it is easy to select an exact number for the amount of clusters. However in random networks it is not so easy to select the optimum number of clusters, it can be found by simulation as shown in Figure 4.2.

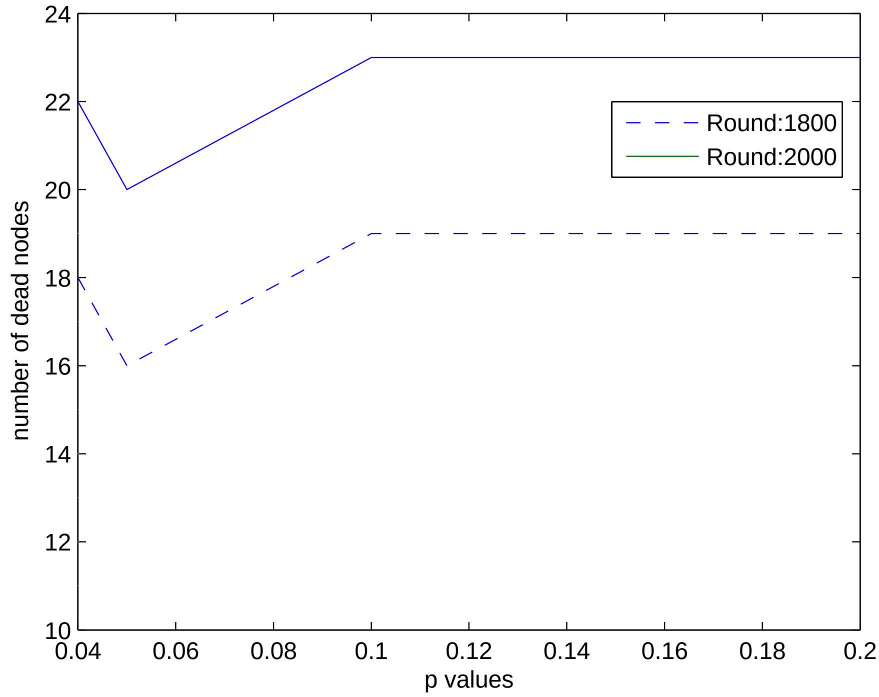

Figure 4.2. Comparison of p values for different rounds.

Smaller values of variance means nodes are located nearly close to each other and the center so there is no need for large number of clusters. If the variance gets larger it means there will be less nodes that are located close each other, nodes will be distributed to whole area. Therefore larger number of clusters is expected to reach more efficient results.

4.4. Clusters and Cluster Heads

LEACH provides an energy effective clustering scheme based on a random clustering. LEACH algorithm does not allow any node become a cluster head more than once in each 1/p rounds. This idea results out a random clustering that improves the results of direct communication as expected. However this random clustering may also result in inefficient clusters and clusterheads. There are different approaches to improve LEACH and solve bottleneck problems of algorithm as the algorithms that are provided in this thesis.

In order to improve LEACH first a deep analyze should be done on algorithm. Due to this fact different simulations are made on LEACH. In order to compare also distribution types, algorihm is applied on two different network topologies.

Figure 4.3, 4.4, 4.5 and 4.6 show the clusters and cluster heads that are selected with LEACH. Different markers are used to show different clusters and cluster heads are marked both with square and its clusters marker. These figures obviously show that LEACH does not always provide robust clustering. In some cases it leads an undesired clustering and this situation may cause an increase in energy consumption. For example in Figure 4.3 cluster heads are not located well , two of the cluster heads are the neighbour nodes.

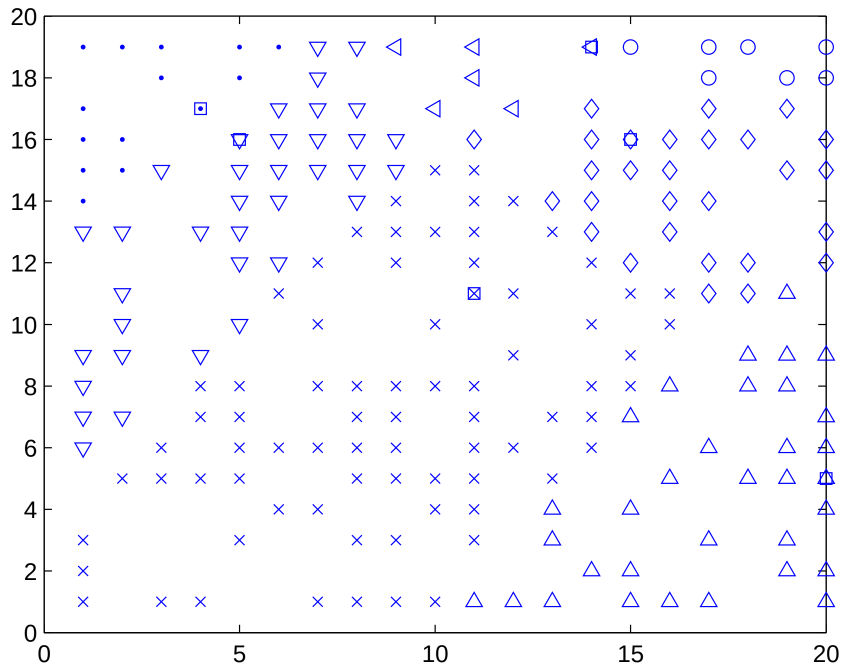

Figure 4.3. Clusters and cluster heads of first round.

The selection of p as 0.05 means that any node can become a cluster head only once in each 20 rounds. Figure 4.7 shows the clusters and cluster heads of 19th round. Since there are not many remaining nodes for cluster heads there occurs less clusters due to less cluster heads. In this scheme clusters have bigger geographical areas and distance within clusters increases. Table 4.2 shows the results that are achieved by the

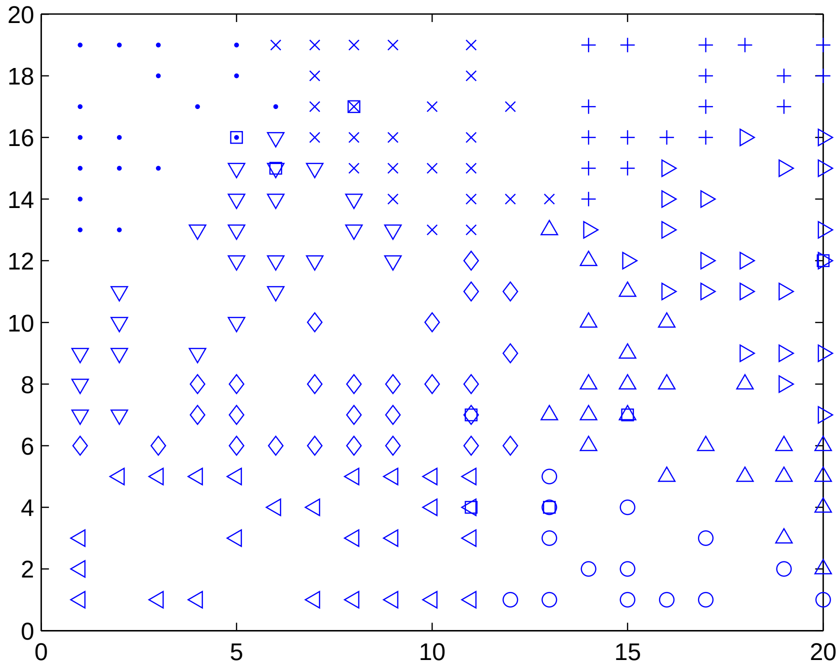

Figure 4.4. Clusters and cluster heads of second round.

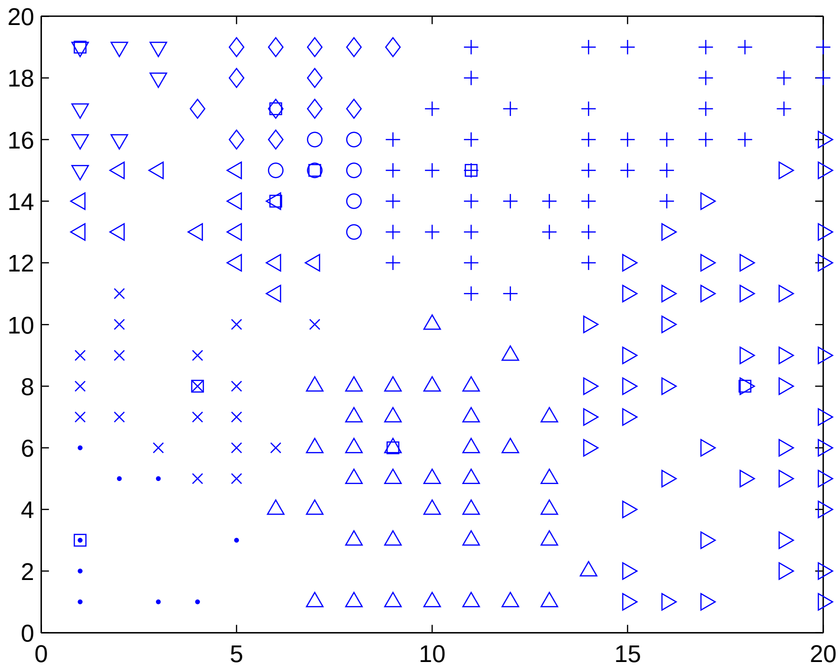

Figure 4.5. Clusters and cluster heads of third round.

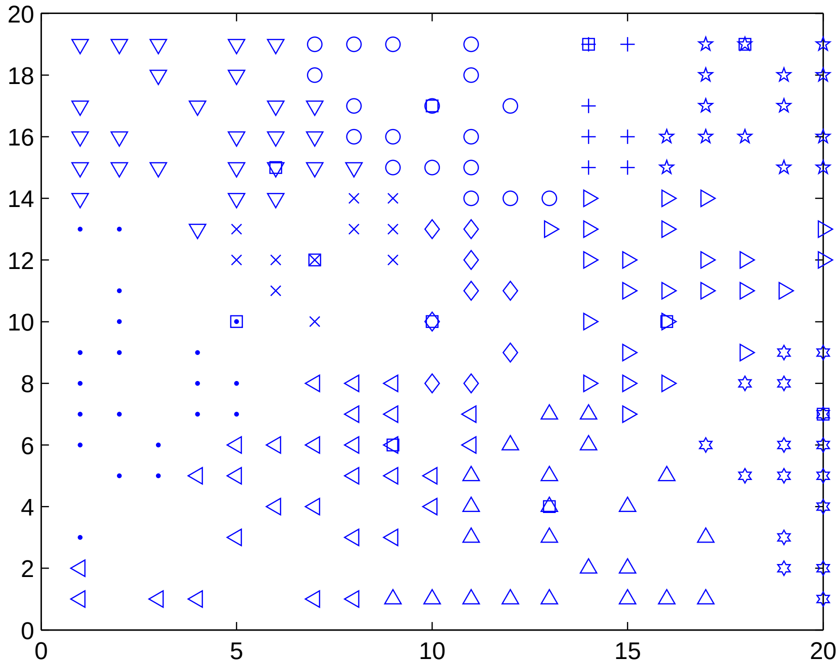

Figure 4.6. Clusters and cluster heads of fifth round.

implemetation of LEACH algorithm. For comparison number of dead nodes after a certain round is used as measurement.

Table 4.2. Simulation results of LEACH.

Round	Number of dead nodes in LEACH	Number of dead nodes in direct way
1600	18	26
1800	28	32
2000	36	38

As it can be seen from Table 4.2 LEACH achives better results than direct communication as expected.

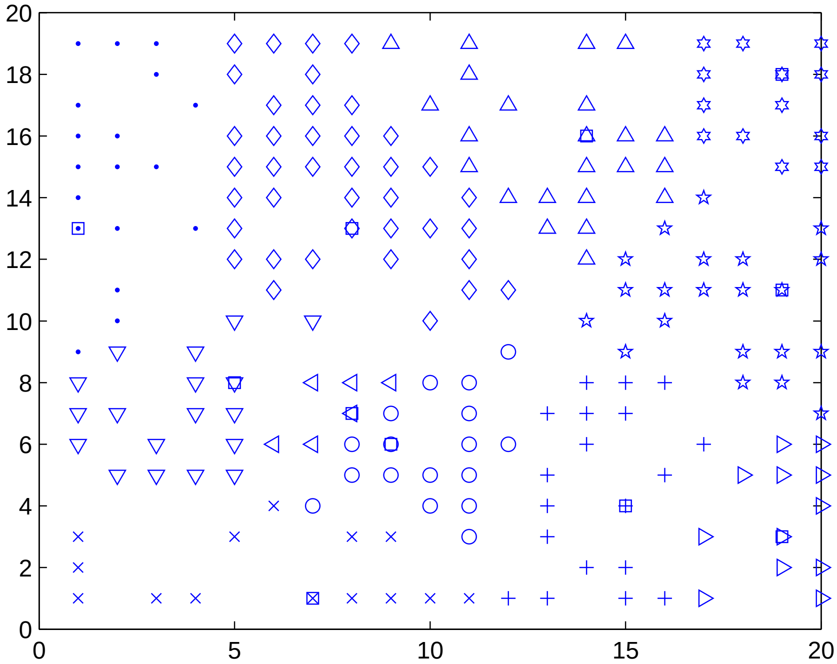

Figure 4.7. Clusters and cluster heads of 19th round.

5. IMPROVED CLUSTERING TECHNIQUE

The main purpose of LEACH based clustering techniques is to provide an energy efficient communication for wireless sensor networks and LEACH is one of the first clustering techniques that is provided for this aim. LEACH has achieved the aimed improvement in many cases but it still is not enough for desired energy optimization. An improved algorithm for network clustering is given in this thesis and main aim of the improvement is to provide a better energy consumption and more efficient communication.

This thesis provides a deep analyze of LEACH and main disadvantages of this algorithm has been tried to be discovered within this analysis. As it can be obviously shown with simulations LEACH algorithm will not always result with a desired situation of clustering therefore a robust optimization in energy consumption is not always expected. In order to improve the efficiency and build a more robust clustering structure we defined four different parameters that are given as connectivity, energy levels of nodes, distance between cluster heads and message transmission rate.

5.1. Cluster Head Selection with Higher Connectivity

Since the sensors in the networks are energy constrained the lifetime of these networks has become a major concern. In order to prolong the networks lifetime cluster heads should be richer in sources than other simple nodes. Therefore intra cluster communication cost will also decrease due to the richer cluster heads. If possible, CHs should be placed close to most of the sensors in its clusters [34],[35]. A new algorithm which aims to select more centralized cluster heads and increase efficiency of the network by decreasing the intra cluster communication costs is given in this section.

As a first step our algorithm differs from LEACH by defining a set of probable cluster heads. LEACH algorithm gives equal chance to all nodes to become cluster head and in this case there may occur some undesired cases such as border nodes becoming cluster heads, or the nodes with no neighbours becomes cluster head. This case of undesired cluster head selection of LEACH is given by Figure 5.1. The cluster marked with + contains so many members but the cluster head selection does not seem to be the optimum one. The cluster head that is marked with square is a border node. Most of the nodes are far away from their cluster head and on the other hand distance between the other cluster heads are not large enough. In this case any communication will cause more energy consumption due to the large distance between cluster heads.

Figure 5.2 shows the second bad case that is occured due to same reason, namely random cluster head selection, in this case the three of the cluster heads are neighbour nodes. Therefore distance between cluster heads is not enough for efficiency.

In order to prevent occurance of these bad cases a set of probable cluster heads is defined according to connectivity levels of nodes. According to the algorithm which is called Algorithm C, the nodes that have at least two neighbours in an area of 2 hops becomes probable cluster heads. By this limitation it is aimed only more centralized nodes to have probability to become a cluster head. Therefore distance between nodes and their cluster heads will be small enough to achieve the aim of energy saving. Figure 5.5 shows a sample case for new algorithm, and as it can be seen from Figure 5.5 the cluster heads are all selected from centralized nodes.

Topology or neighbour discovery in sensor networks is generally done by letting nodes send hello messages in order to signal their presence [36]. In order to apply new connectivity algorithm nodes need to discover number of their neighbours and to do this nodes can send hello messages to all neighbors within a predetermined number of hops. Each node can count the number of hello messages it receives. After this process is completed connectivity algorithm can be applied to the network. Within the application of this algorithm there will be a neighbour discovery process at the beginning but on the other hand after the topology discovery is completed number of probable cluster heads will be determined. The nodes that does not have enough degree of connectivity will not have a chance to become a cluster head and they will

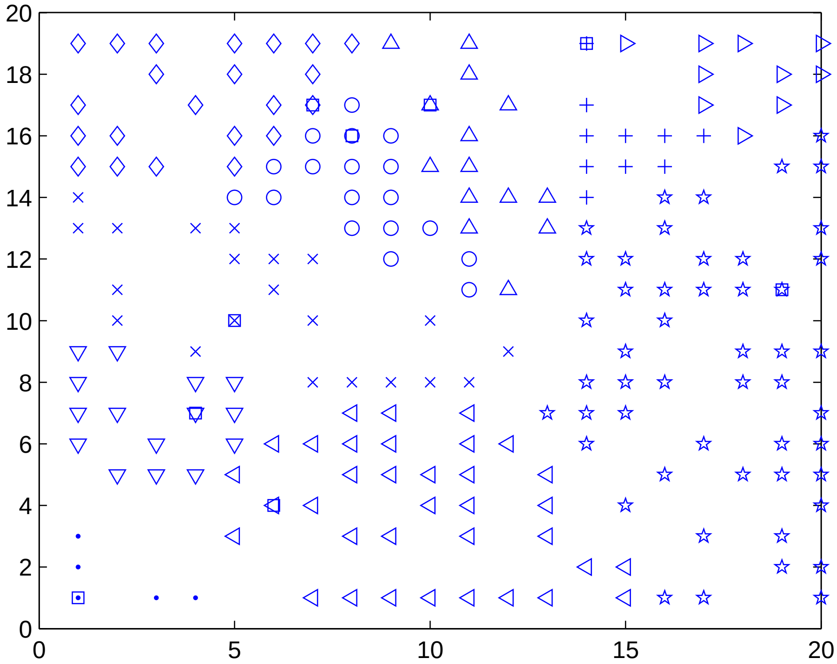

Figure 5.1. Bad clustering case for LEACH-border node.

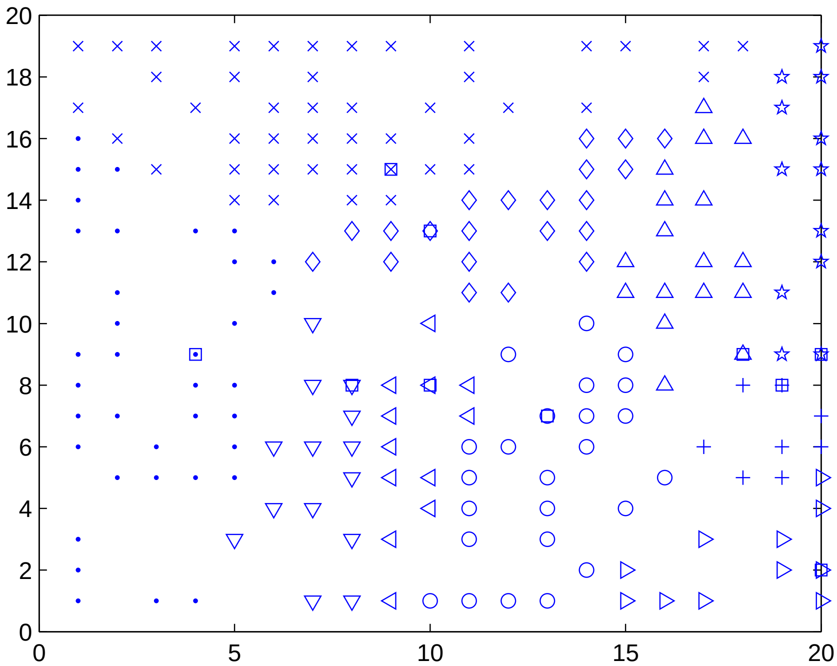

Figure 5.2. Bad case of LEACH-neighbour cluster heads.

not need to run the cluster head selection algorithm in any round. Only process that will be done by these nodes will be to find their appropriate clusters. This will reduce the overall process load. Flow chart of connectivity algorithm is given in Figure 5.3.

Simulation results show that applying the connectivity parameter to LEACH improves the efficiency of the network. Table 5.1 shows the improved results.

An analysis of Figure 5.4 shows connectivity parameter has a better effect on networks lifetime compared to LEACH. Applying the new Algorithm C increases the efficiency and lifetime of network.

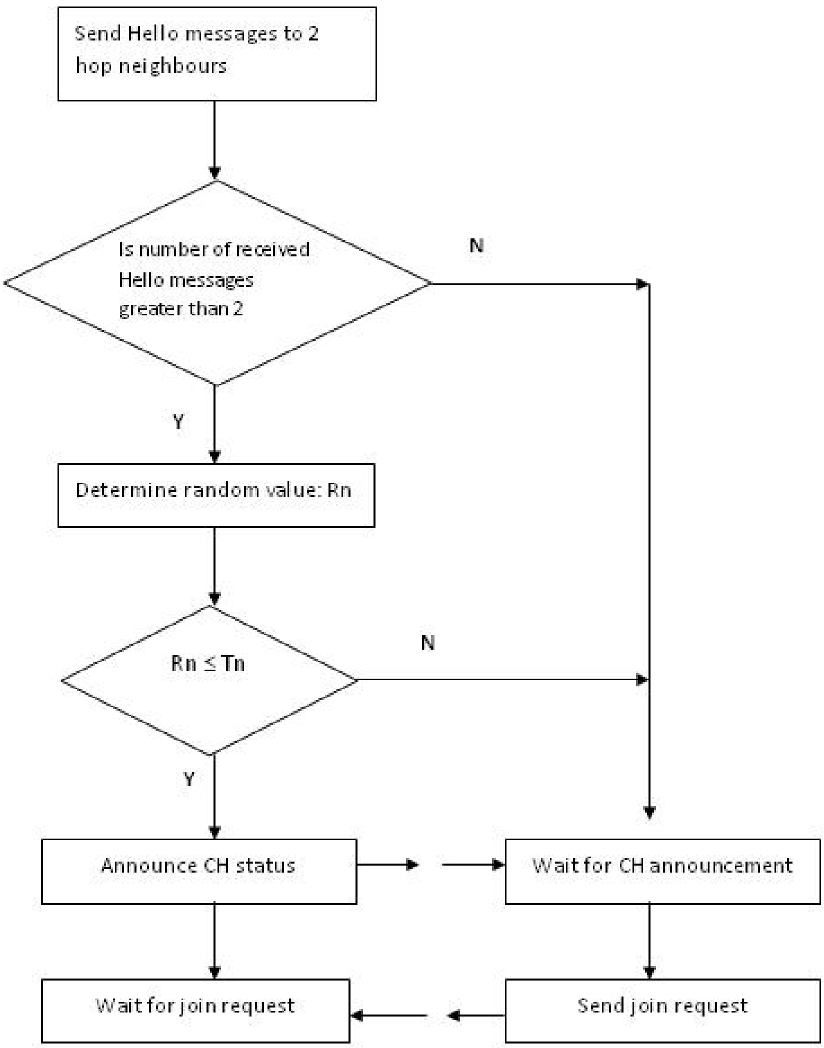

Figure 5.3. Flow Chart of Connectivity Algorithm.

Table 5.1. Connectivity results.

Method	Round	Percentage of dead nodes
LEACH	1200	5,5
Algorithm C	1200	5
LEACH	1400	8
Algorithm C	1400	7,5
LEACH	1600	10,5
Algorithm C	1600	10
LEACH	1800	14,5
Algorithm C	1800	13,5
LEACH	2000	19
Algorithm C	2000	17

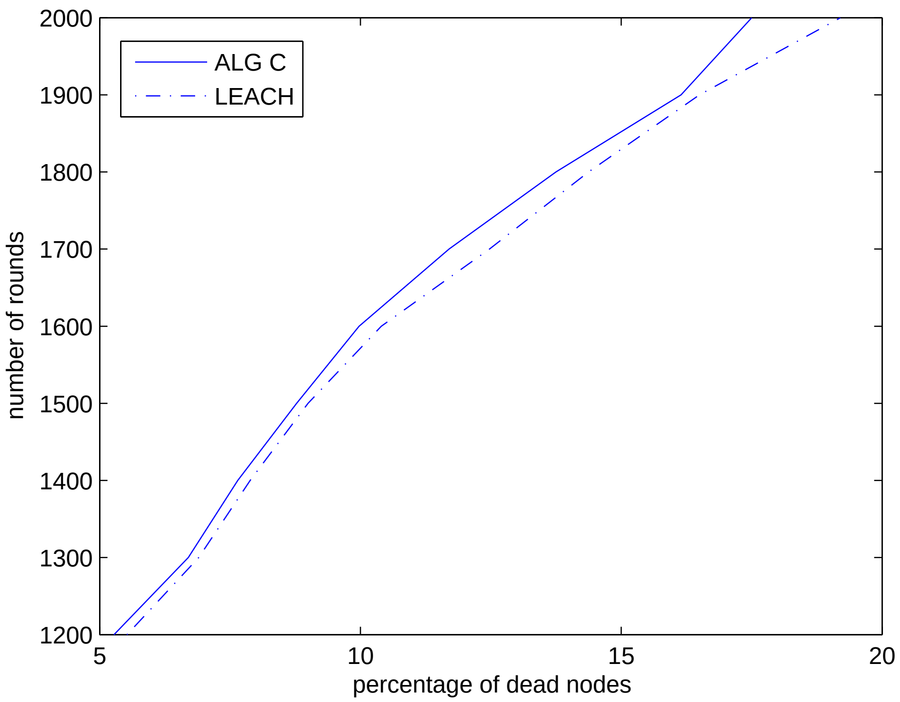

Figure 5.4. Simulation results for connectivity algorithm.

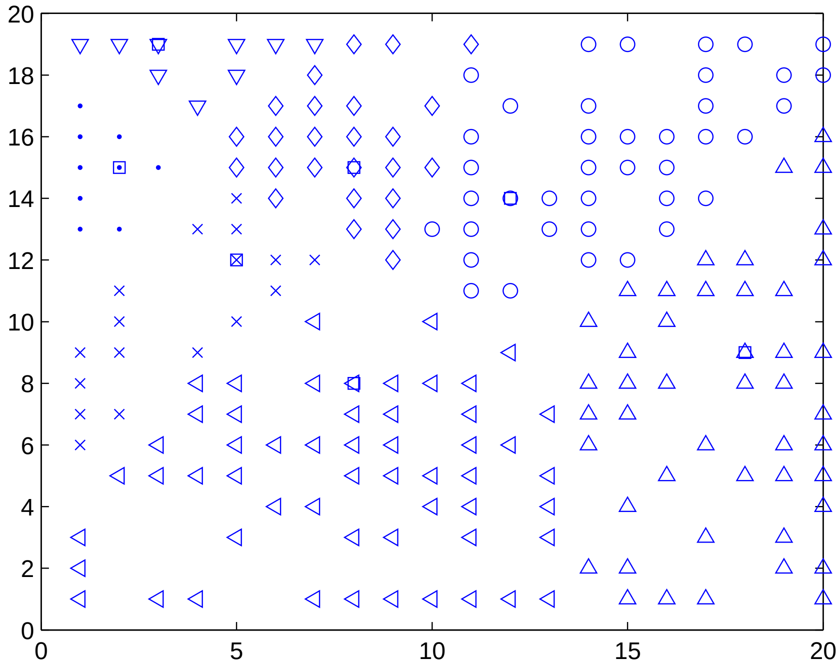

Figure 5.5. A sample clustering case for the Algorithm C.

5.2. Energy Constrained Random Clustering

By defining a set of probable cluster heads and selecting the cluster heads from this set it is expected to cause better results in case of energy saving, however there are still some undesired clustering cases that can occur with this algorithm. One of the most important points in wireless networks is the energy levels of nodes. Because the communication based on clustering is directly related with energy levels of nodes.

Since all the communication between different clusters are provided via cluster heads, these nodes becomes the ones that require more energy than other simple cluster members. If the energy level of cluster head nodes are not adequate enough to provide this communication it will become a bottleneck for the network. Selection of the cluster head from a set of probable cluster heads will cause less nodes to have probability to become a cluster head. LEACH aims to provide energy optimization by adaptation of cluster heads but the only method used in LEACH is based on random functions.

In the Algorithm C it is aimed to select cluster heads from more centralized nodes but it may cause network to become stuck in higher rounds. Because these centralized nodes will lose more energy than other nodes. Since our main focus is to optimize the energy consumption we need to consider energy levels of the cluster head nodes. Therefore one of the parameters that will be used in cluster head selection becomes the definition of energy levels of probable cluster heads. In each round energy levels of the members of probable cluster head set are calculated, and this value is compared with a predefined threshold. If nodes have enouh energy to become a cluster head they will be evaluated in cluster head selection, in other case these nodes will leave the set of probable cluster heads.

Initial energy levels of sensors can vary due to the different characteristics and types of equipments. The initial energy levels of most commonly used sensors in wireless sensor networks vary between 0.1J and 2J. We assumed that all the nodes have same level of energy that is equal to 0.5J [25] in the beginning and in cluster head selection, threshold value of energy for becoming a cluster head is defined as 0.1J. Therefore the nodes which lose 80 percent of their initial energy will not have chance to become a cluster head anymore. The flow chart for this algorithm which is named as Algorithm E is given in Figure 5.6.

The threshold value for enegy level becomes very important for the larger rounds. Higher values of threshold may have good results in the smaller rounds however if this threshold value is very large in the higher rounds network may become stuck. Because the higher energy level nodes will become cluster head and they will lose some amount of their energy in the rounds which they act as cluster heads. Therefore in the higher rounds it is expected all the nodes to lose some amount of their energy and there will be no higher energy level node. If the threshold value for cluster head selection is high it will cause network to become stuck in the last rounds. On the other hand smaller values of this threshold will be meaningless in Algorithm E. Since our aim is to determine higher energy level nodes as the probable cluster heads the selection of a small threshold value is not expected to improve the results.

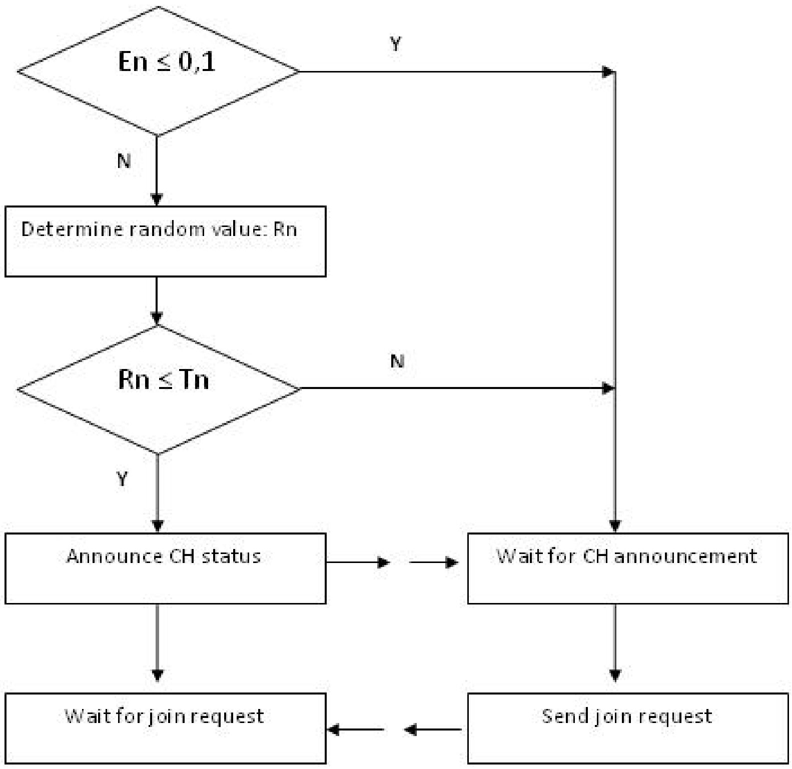

Figure 5.6. Flow Chart of Algorithm E.

There are different algorithms in the literature as the improved versions of LEACH by energy factor [5],[26],[27]. Some of the examples of these algorithms are given in Section 2.1. All of these algorithms are based on the assumption that each node can calculate its remaining energy level. According to these algorithms nodes need to calculate their energy levels in each round. The authors in [27] provided a new threshold value in order to select higher energy level nodes that is decreased by the factor of energy levels.Therefore nodes will need to calculate their current energy level in each round in order to determine the threshold value for that round. In [5] each node computes the quotient of its own energy level and the aggregate energy remaining in the network. With this value each node decides if it becomes cluster-head for this round or not. High-energy nodes will more likely become cluster-heads than low-energy nodes. The disadvantage of this approach is, that each node has to estimate the aggregate remaining energy in the network since this requires additional communication with the base station and other nodes. The process of our new algorithm is similar with the older ones. Nodes need to calculate their energy levels in each round until they have less energy than the threshold.

It can be obviously seen from simulation results that considering the energy levels of the nodes enhances the clustering algorithm. Table 5.2 shows the simulation results when energy level parameter is applied to LEACH algorithm.

Figure 5.7 shows the improvement of Algorithm E on LEACH and it can be obviously seen from these results that applying energy parameter to the LEACH gives better results.

5.3. Elimination of Close Cluster Heads

After defining prerules of cluster head selection algorithm, an analyze has been done on the new clustering technique. There are still some drawbacks of this clustering scheme. Although all nodes do not have the probability to become a cluster head, cluster head selection and location of these cluster heads does not always result in a desired topology. The analysis of Figure 5.8 obviously shows that cluster head nodes

Table 5.2. Simulation results of energy algorithm.

Method	Round	Percentage of dead nodes
LEACH	1200	5,5
Algorithm E	1200	3,5
LEACH	1400	8
Algorithm E	1400	6
LEACH	1600	10,5
Algorithm C	1600	10
LEACH	1800	14,5
Algorithm E	1800	8,5
LEACH	2000	19
Algorithm E	2000	12,5

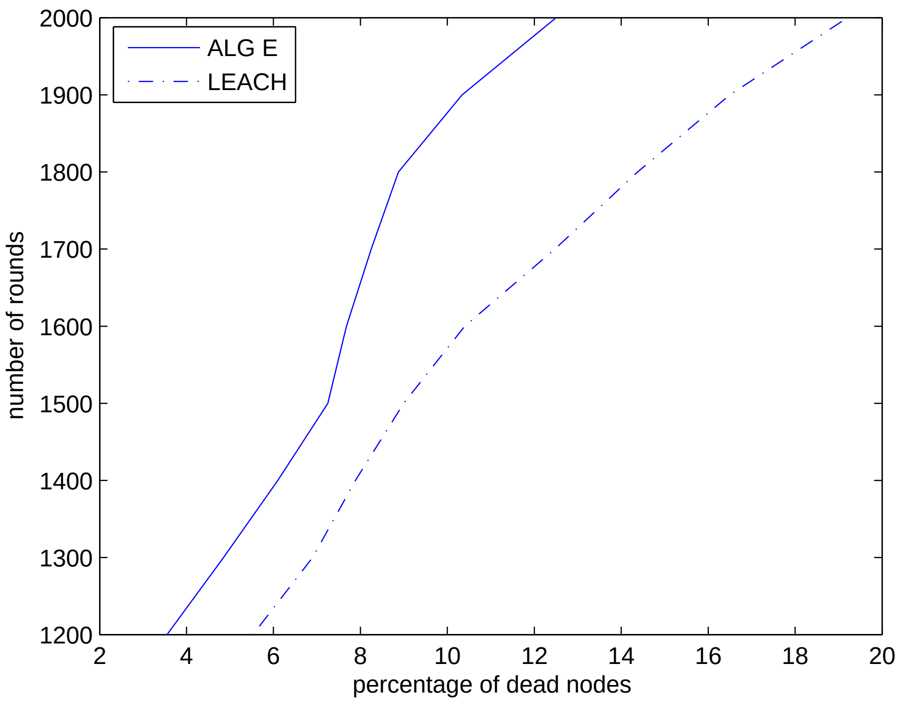

Figure 5.7. Simulation results of energy levels algorithm.

are located in central positions and they have more than 2 neighbours but the distance between the cluster heads are too small. Most of the cluster heads are neighbour nodes. Therefore this structure leads an undesired case in case of efficiency. For example in this kind of cluster head occurance two neighbour nodes can be members of different clusters and therefore they cannot send message to each other directly. Due to the main idea of hierarchical routing transmitter node first needs to send the message to its cluster head and a transmission between these two cluster heads will begin and finally receiver will receive the message from its own cluster heads. Therefore this transmission will include four hops instead of two hops which would occur if these two neighbour nodes belong to same cluster.

There may be different solutions to the problem of neighbour cluster heads one of these solutions can be integration of these two groups. However in this case we need to eliminate one of the cluster heads and select the other one as the cluster head of merged clusters. The basic of this elimination has also an important role in clustering method, it will shape the new clustered network. Since the main aim of this clustering algorithm is energy optimization, the energy levels of these two nodes is decided as a measure for elimination. The cluster head with higher energy will remain as cluster head and the other one will be eliminated.

After this elimination these two clusters can merge and a new cluster may occur with the higher energy cluster head. Although the problem of neighbour cluster heads can be solved by this elimination method the integration of two clusters will result out a bigger cluster. Therefore intra-cluster communication may cause more energy dissipation. A new method is provided in order to prevent the occurance of larger clusters during the elimination of neighbour cluster heads. Before making the clustering first cluster heads are determined and elimination of the neighbour cluster heads will be done in this phase. When the cluster heads are determined an algorithm is applied to these cluster heads. Distance between these selected cluster heads are calculated and the existince of any neighbour cluster heads is determined. If any neighbour cluster heads are found the elimination formula is applied to these nodes. Therefore one of them will be selected as a cluster head and the other one will remain as a simple node

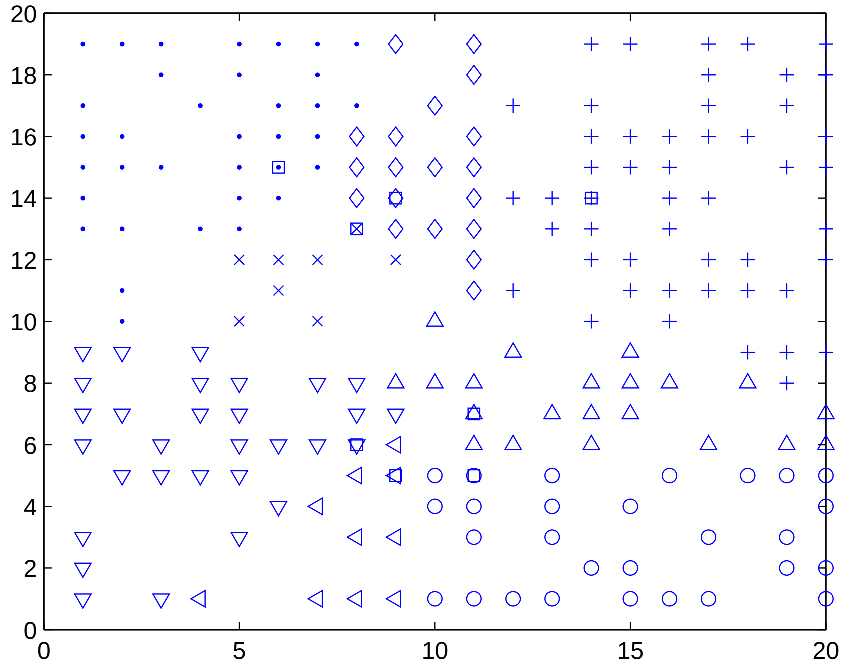

Figure 5.8. Neighbour cluster heads.

according to their energy level.

After cluster head selection phase the clusters will be determined according to clustering algorithm. Every node will calculate their distances to cluster heads and they will select the closest one as their cluster head. As a result distance between all the cluster heads will not be small enough to cause undesired cases and probabilty of larger clusters will decrease. Due to the usage of the parameter distance between cluster heads new algorithm is named as Algorithm D. Figure 5.9 shows the sample case of clustering that is occured with this algorithm. When the elimination of close cluster heads is applied to algorithm simulation results show that it has achieved an important improvement on networks efficiency. Table 5.3 shows the simulation results of new algorithm.

It can be obviously seen from Figure 5.10 that Algorithm D achieves a big improvement on LEACH. Since this algorithm reshapes the clustering structure and prevents the undesired clustering cases it results out a more efficient network. Since

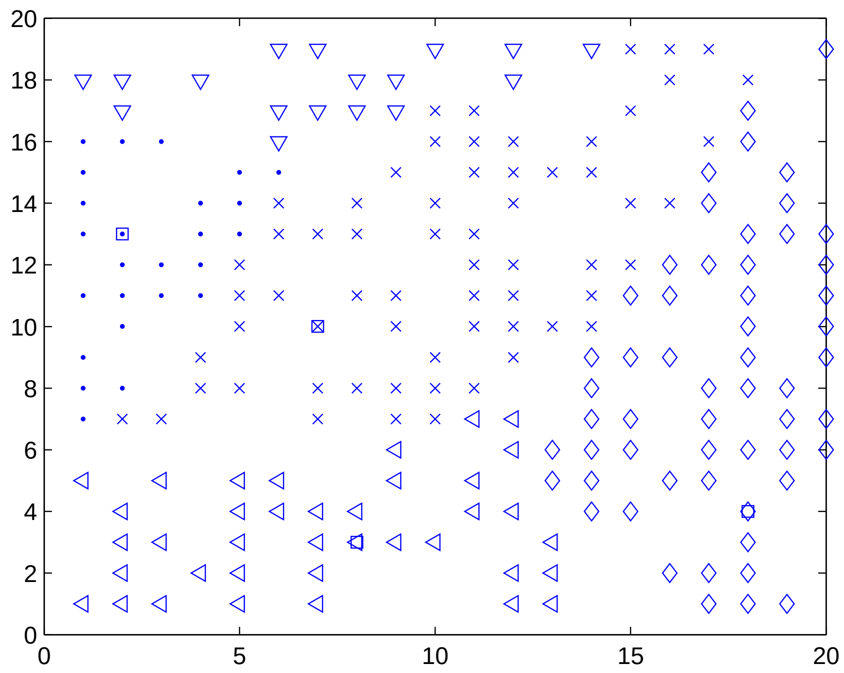

Figure 5.9. Simple clustering case with Algorithm D.

Table 5.3. Simulation results for close cluster head elimination.

Method	Round	Percentage of dead nodes
LEACH	1200	5,5
Algorithm D	1200	4
LEACH	1400	8
Algorithm D	1400	6
LEACH	1600	10,5
Algorithm D	1600	8
LEACH	1800	14,5
Algorithm D	1800	11,5
LEACH	2000	19
Algorithm D	2000	16

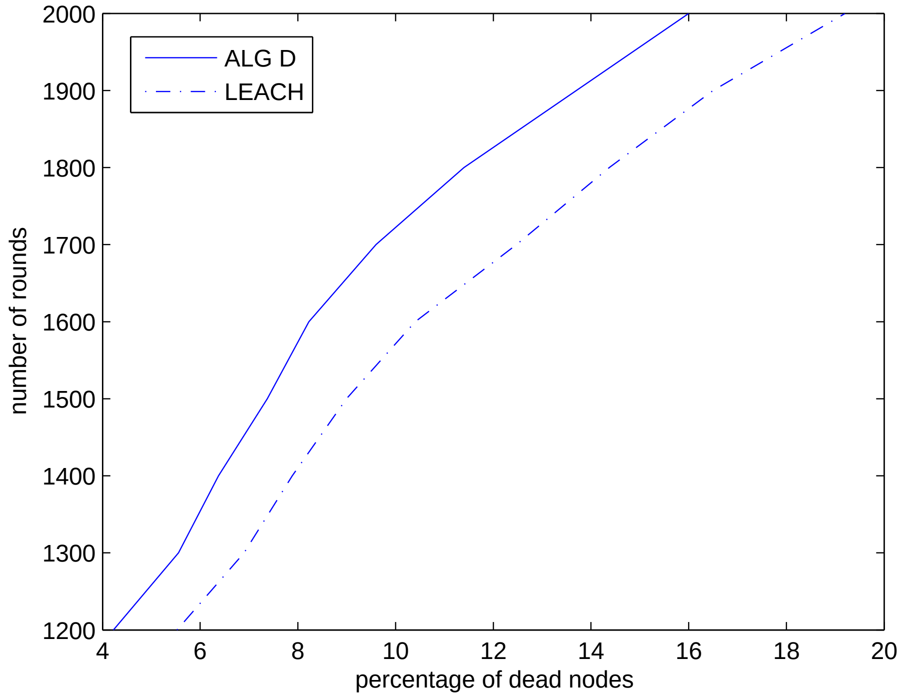

Figure 5.10. Simulation results of algorithm D on LEACH.

LEACH has a random based clustering type it cannot prevent undesired clustering cases to happen and all the other parameters mentioned in this thesis except distance just selects some nodes to have more probability to become cluster head. Even some nodes are selected to have more probability some undesired clustering cases can still occur. Although some constraints are applied to the network these parameters does not reshape the clusters. However Algorithm D checkes the clustered network and cancels the clustering if needed. Moreover in case of reshaping we used the energy levels for comparison and this is expected to increase the effect of improvement.

In the proposed algorithm focus was placed on the check and correction of the random clustering. However in order to apply this algorithm we need to know the locations of the nodes. Since these sensors are located on larger geographical areas and usually placed randomly it is too difficult to have the knowledge of locations. On the other hand some researchers [37],[38] assumed that the global position system (GPS) [39] was installed in each sensor node. However, high costs and high power-consumptions of GPS are heavy burdens for sensor nodes. Furthermore, due to the interference of buildings and terrain obstructions, GPS can not work well in scenes such as the indoor, the seabed, and the battlefield [40]. Instead of using GPS nodes can send Hello messages to a predefined number of hop distance in order to discover neighbour cluster heads. Once the cluster heads are determined these nodes will send Hello messages to a distance of 2 hops. This is a small message containing Node's ID and energy level. If a cluster head node receives a Hello packet it will check the energy information that is included in Hello message. If the node's own energy level is higher than the energy level in the message than this node will remain as cluster head. Otherwise the node will not stay as cluster head and will start to wait for cluster head announcements. Flow chart for this algorithm is given in Figure 5.11.

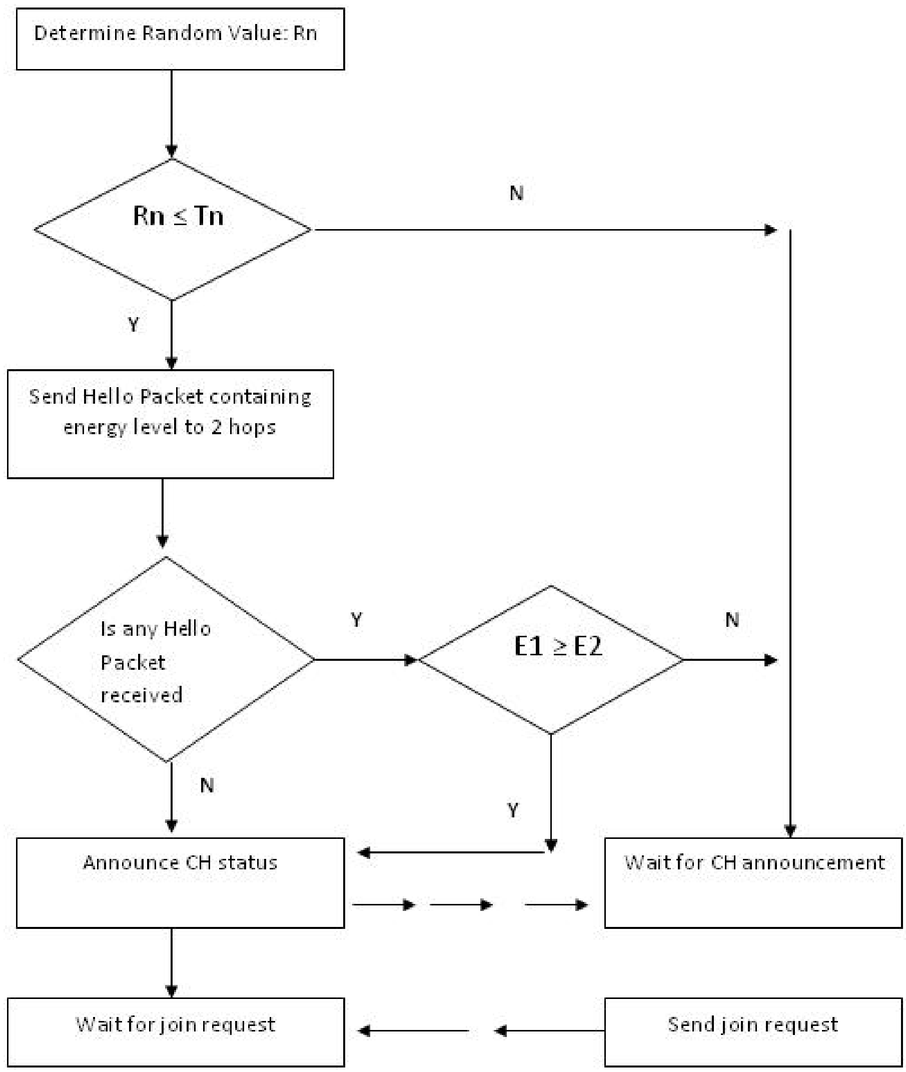

Figure 5.11. Flow Chart for Algorithm D.

5.4. Priority in Cluster Head Selection with Transmission Rate

Finally a new challange and improvement is provided with taking into account the message receive and transmit rates of nodes in cluster head selection. The main idea behind this improvement is if a node is more eager to send or receive messages it should have more probability to become a cluster head. In this case it will not need to use more energy in every message transmission. Since we applied the parameter of message transmission rate the algorithm is named as Algorithm R.

In order to reach this improvement it is assumed that a known communication scenario is applied to network. Since cluster heads are selected in each round, this rate of transmission willingness is taken into account by modifiying the random values that are defined by each node.

If a node has more role in communication scenario even as a receiver or a sender the random value of this node will decrease by a factor of its message transmission rate. The new random value will be modified as in Formula 5.1. The parameter MRT stands for the message transmission rate of a node which is calculated before the clustering. Since we assume that a known communication scenario is applied, the roles of the nodes in this communication scenario can be discovered before clustering. Therefore message transmission rate can be configured priori at each node. The flow chart of Algorithm R is given in Figure 5.12.

$$R_{n_{new}} = \frac{R_n}{MRT} \tag{5.1}$$

5.5. Comparison and Combination of New Clustering Algorithms

As it is given in previous sections four new parameters are provided to improve the LEACH clustering algorithm such as connectivity, energy levels of nodes, distance

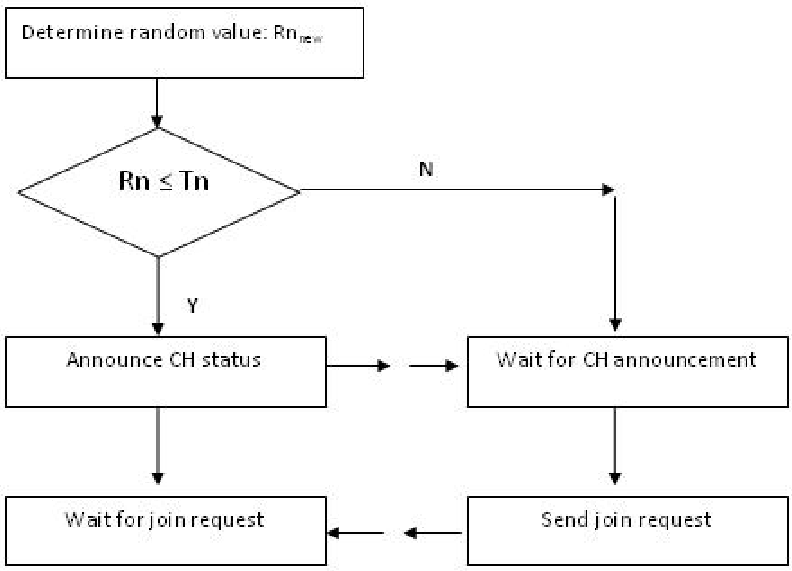

Figure 5.12. Flow Chart for Algorithm R.

Table 5.4. Simulation results of algorithm R.

Method	Round	Percentage of dead nodes
LEACH	1200	5,5
Algorithm R	1200	5
LEACH	1400	8
Algorithm R	1400	7,5
LEACH	1600	10,5
Algorithm R	1600	10
LEACH	1800	14,5
Algorithm R	1800	13,5
LEACH	2000	19
Algorithm R	2000	18

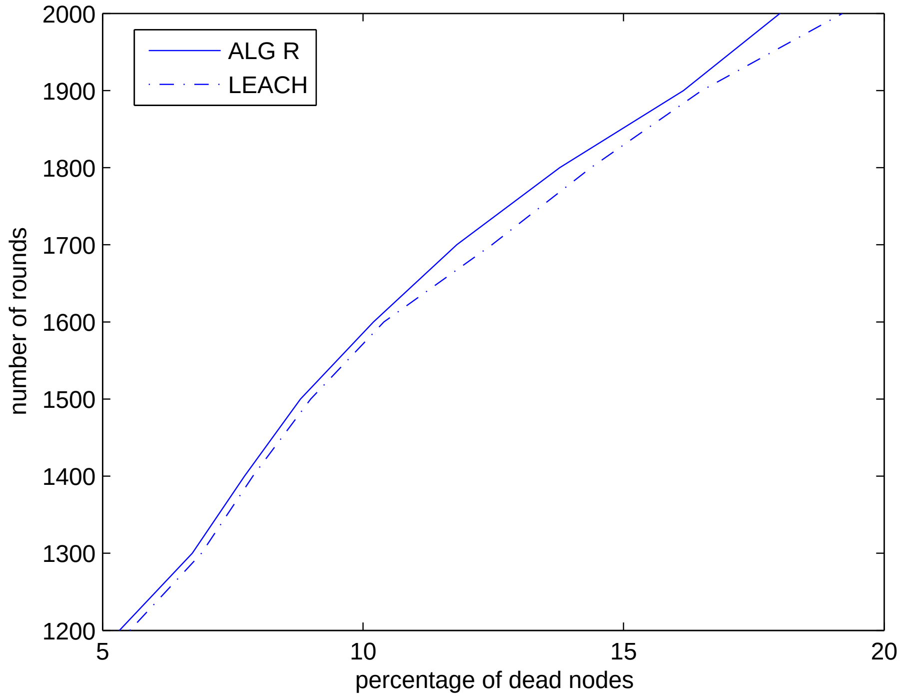

Figure 5.13. Simulation results of algorithm R.

between cluster heads and transmission rate. Simulation results show that all of these four parameters improved the algorithm indiviually and as the final step it is planned to combine and compare all these parameters to provide the final enhanced clustering algorithm.

In order to compare new methods with LEACH and measure their effectiveness in case of energy optimization, same simulations have been done both on LEACH and new algorithms.

As the first step of simulation the new algorithm that is improved by connectivity and energy levels is applied to the network.For all simulation cases it is assumed that all nodes have 0.5J initial energy and same communication scenario is applied for all cases. In order to compare two clustering algorithms number of dead nodes is used as the measurement parameter. Table 5.5 shows the simulation results for algorihm CE.

The analysis of Table 5.5 shows that combination of these two parameters has

Table 5.5. Simulation results of algorithm CE.

Method	Round	Percentage of dead nodes
LEACH	1200	5,5
Algorithm CE	1200	3,5
LEACH	1400	8
Algorithm CE	1400	6
LEACH	1600	10,5
Algorithm CE	1600	7,5
LEACH	1800	14,5
Algorithm CE	1800	8,5
LEACH	2000	19
Algorithm CE	2000	12,5

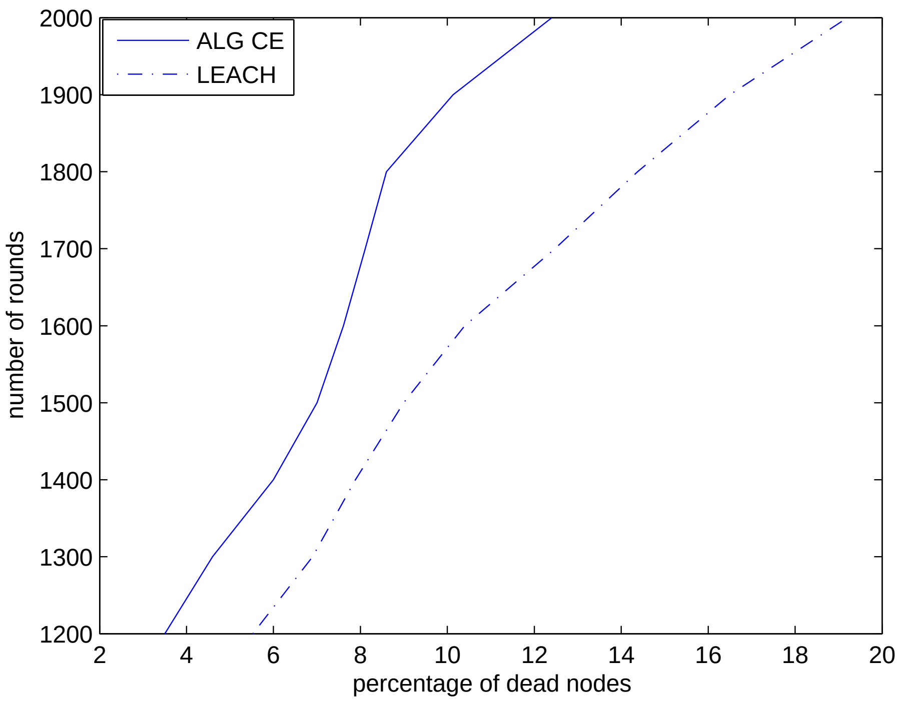

Figure 5.14. Simulation results of algorithm CE on LEACH.

also enhanced the clustering algorithm. Analysis of Figure 5.14 shows that combination of these two algorithms makes the network more efficient. While connectivity selects the more centralized nodes to be member of probable cluster heads energy parameter prevents low energy level nodes to become cluster head. Therefore combination of these two algorithms leads a better algorithm.

In order to achieve a better algorithm distance parameter is also applied to Algorithm C. While algorithm C reaches better results than LEACH algorithm DC will check the clustering and reshapes the clustered network. Therefore a bigger improvement is expected to achieve with algorithm DC.

Table 5.6. Simulation results of algorithm DC.

Method	Round	Percentage of dead nodes
LEACH	1200	5,5
Algorithm DC	1200	4
LEACH	1400	8
Algorithm DC	1400	6,5
LEACH	1600	10,5
Algorithm DC	1600	8
LEACH	1800	14,5
Algorithm DC	1800	11
LEACH	2000	19
Algorithm DC	2000	15

Table 5.6 and Figure 5.15 shows the results of algorithm DC. As the next step energy parameter is applied to algorithm DC. By this application it is expected to prevent low energy level nodes to become cluster head and improve the results of networks lifetime.

Table 5.7 and Figure 5.16 shows the simulation results of algorithm DCE. It can be obviously seen from Figure 5.16 that algorithm DCE has a very big improvement

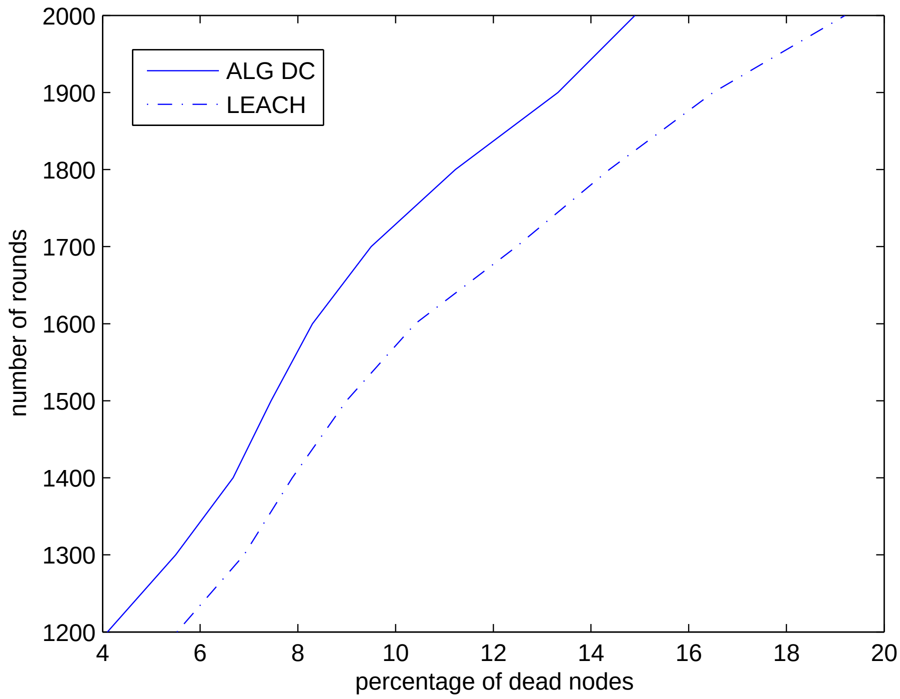

Figure 5.15. Simulation results of algorithm DC on LEACH.

on LEACH when compared with other algorithms.

As the next step message transmission rate is applied to the Algorithm DCE, however results does not seem to be improved when compared with algorithm DCE. An analysis of Figure 5.17 and 5.18 shows that algorithm DCER has also an improvement on LEACH however the results of algorithm DCE is better than DCER. The simulation results can also be seen in Table 5.8.

The analysis of the results for algorithm R and DCER shows that message transmission rate provides an improvement on networks lifetime however it decreases the improvement of algorithm DCE. When algorithm DCE is applied to a network the number of nodes that can become cluster head mainly decreases in comparison with LEACH. Since the main aim of the LEACH clustering is to provide energy optimization with adaptive clustering, it needs more nodes to become cluster head to share the load of transmission. With the algorithm CE we decrease the number of nodes that have chance to become cluster head and when algorithm R is applied it gives priority

Table 5.7. Simulation results of algorithm DCE.

Method	Round	Percentage of dead nodes
LEACH	1200	5,5
Algorithm DCE	1200	2,5
LEACH	1400	8
Algorithm DCE	1400	5
LEACH	1600	10,5
Algorithm DCE	1600	6,5
LEACH	1800	14,5
Algorithm DCE	1800	7,5
LEACH	2000	19
Algorithm DCE	2000	9

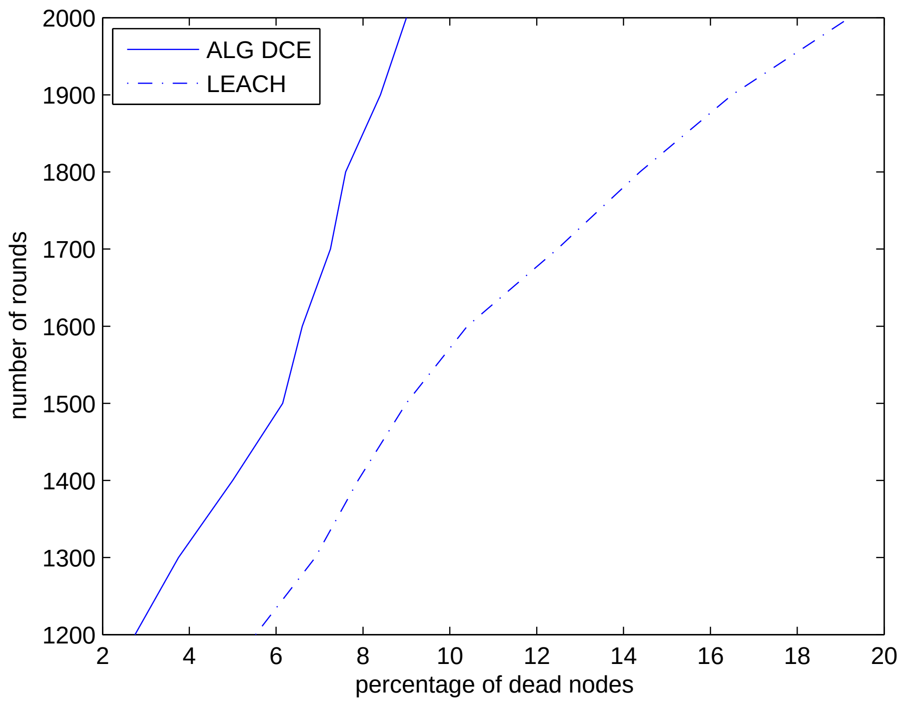

Figure 5.16. Simulation results of algorithm DCE on LEACH.

Table 5.8. Simulation results of algorithm DCER.

Method	Round	Percentage of dead nodes
LEACH	1200	5,5
Algorithm DCER	1200	3
LEACH	1400	8
Algorithm DCER	1400	5
LEACH	1600	10,5
Algorithm DCER	1600	7
LEACH	1800	14,5
Algorithm DCER	1800	8
LEACH	2000	19
Algorithm DCER	2000	9,5

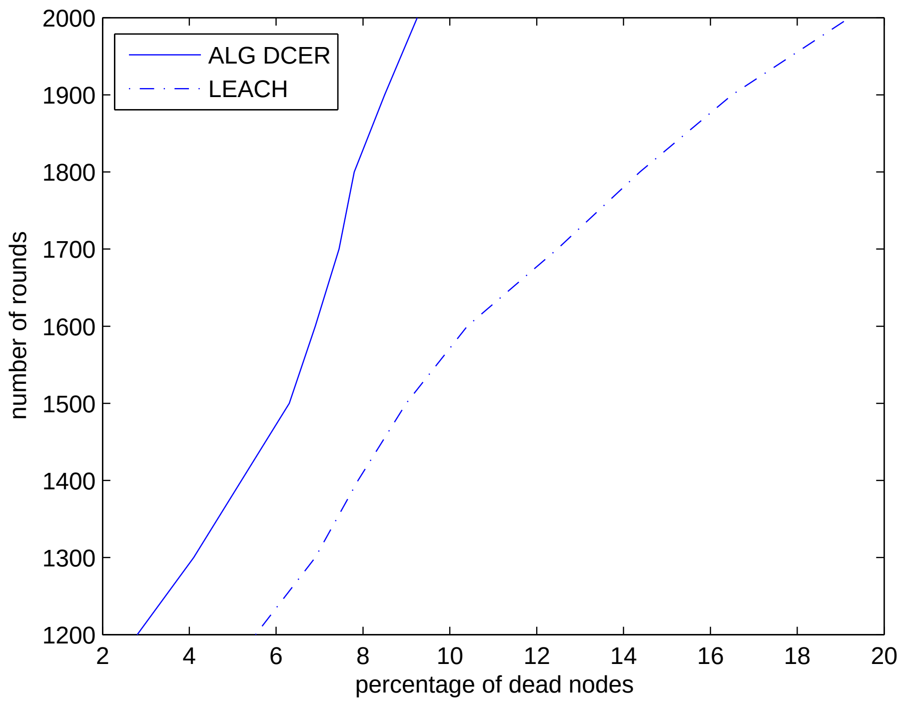

Figure 5.17. Simulation results of algorithm DCER on LEACH.

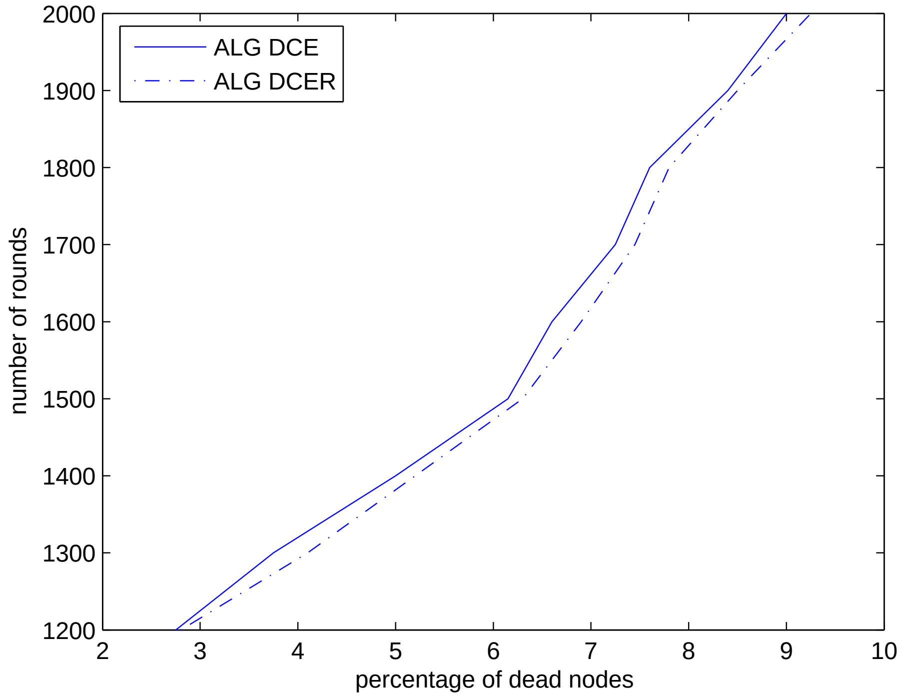

Figure 5.18. Comparison of algorithms DCE and DCER.

to some nodes. Therefore the main aim of the LEACH based algorithm will not be provided efficiently. Transmission load needs to be distributed to all probable nodes equally to get better results because there are not so many nodes that have chance to become cluster head.

Figure 5.19 shows the results of the all algorithms that mentioned in this thesis.

Analysis of Figure 5.19 gives that the algorithm DCE provides the most efficient results for both of the networks.

The analysis also shows that our algorithm DCE has an improvement of networks lifetime in all rounds . As the number of rounds increase improvement on the number of alive nodes also increases until the network becomes stuck. An improvement of 10 percent has been achieved for 2000 rounds.

These results show that number of dead nodes after a certain round will decrease

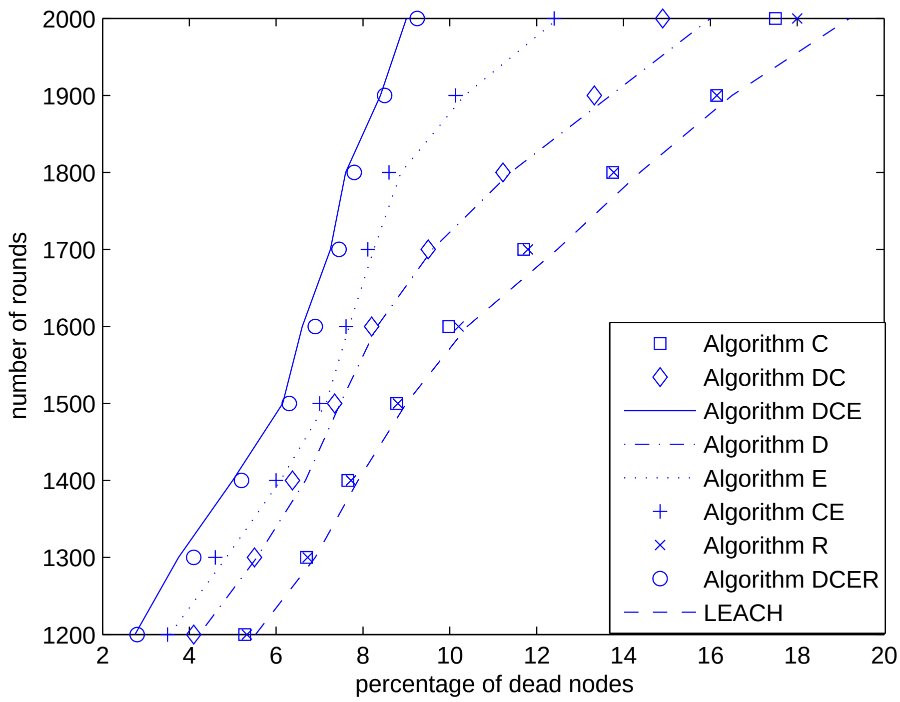

Figure 5.19. Comparison of clustering algorithms.

by applying our new improved algorithm DCE. Figure 5.20 and Figure 5.21 shows the clustering result of a single round for the random network. It can be seen from these figures that both LEACH and new algorithm gives efficent results for clustering however there occurs some undesired cases in LEACH such as close heads. If Figure 5.20 is analyzed it can be seen that with new algorithm cluster heads are usually located in the center of their clusters and this is expected to lead better results in communication phase.

As the communication scenario holds on and the rounds increase the energy level of the nodes decreases. Our main aim is to optimize energy consumption of whole network and prolong the networks lifetime. In LEACH algorithm every node has a probability to become a cluster head but when a node becomes a cluster head for a single round it will cause this node to use more energy. Therefore it is desired cluster heads to have more energy than the other ones in order to provide contionus communication. If the cluster head becomes dead it means communication will be impossible for any node within that cluster for that round. Altough LEACH is improved with new algorithm

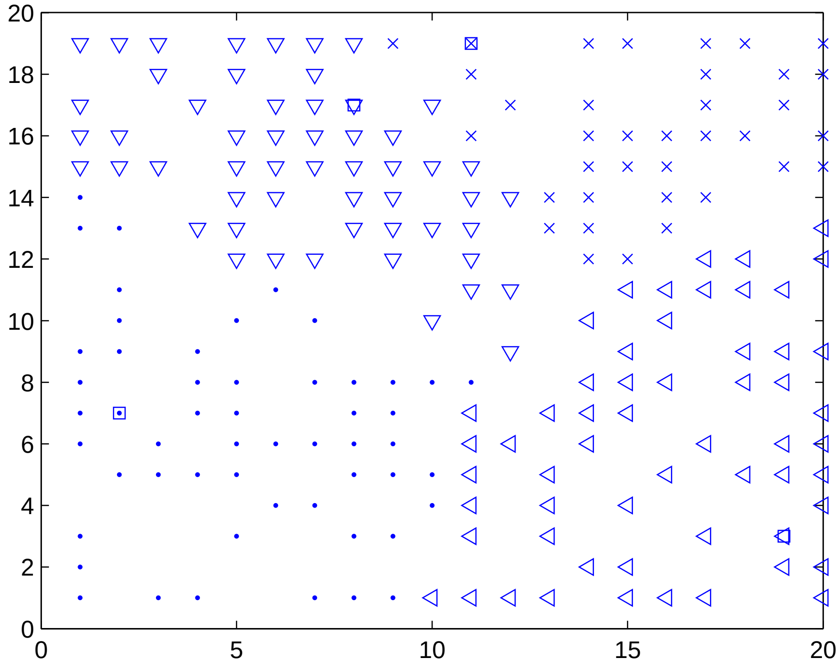

Figure 5.20. Clustering result for 19th round with LEACH.

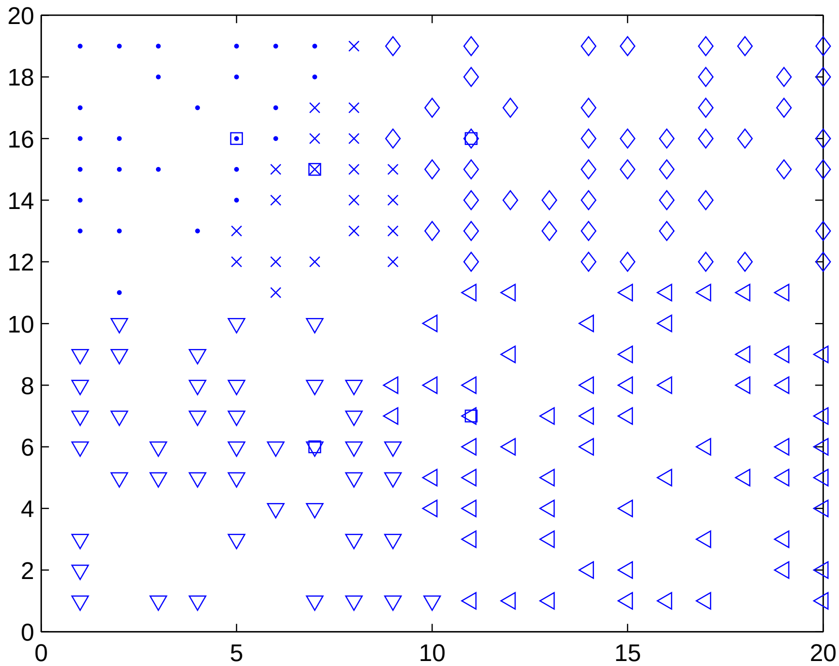

Figure 5.21. Clustering result for 19th round algorithm DCE.

to lead higher energy level nodes to become cluster head, this new algorithm also may make network stuck in higher rounds.

In order to compare cluster head selection in higher rounds we simulated clustering and analyze clusters and cluster heads for higher rounds. In the last part the clustering results for higher rounds are given with Figure 5.22, 5.23, 5.24, 5.25, 5.26 and 5.27.

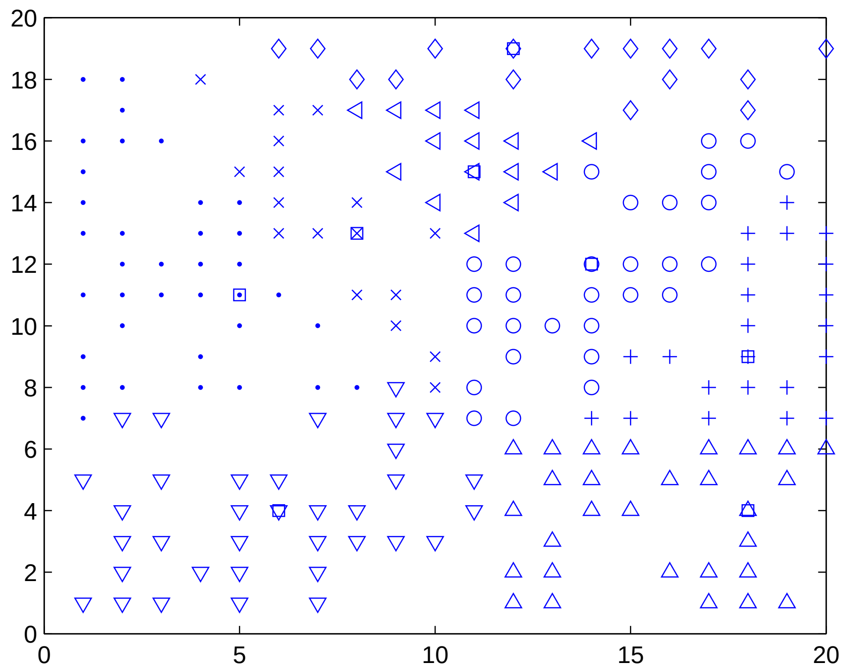

Figure 5.22. Clustering results for round 1200 with LEACH.

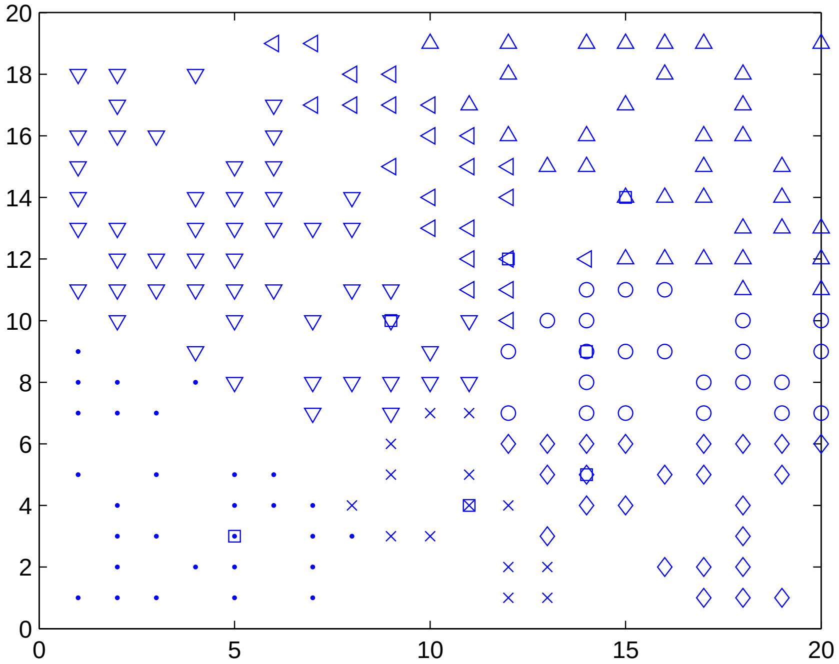

Figure 5.23. Clustering results for round 1200 algorithm DCE.

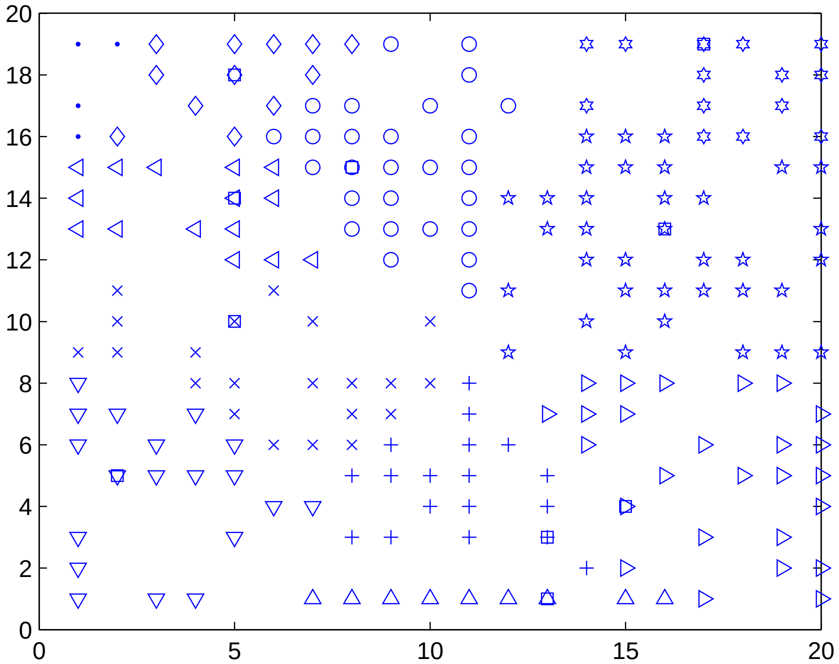

Figure 5.24. Clustering results for round 1500 with LEACH.

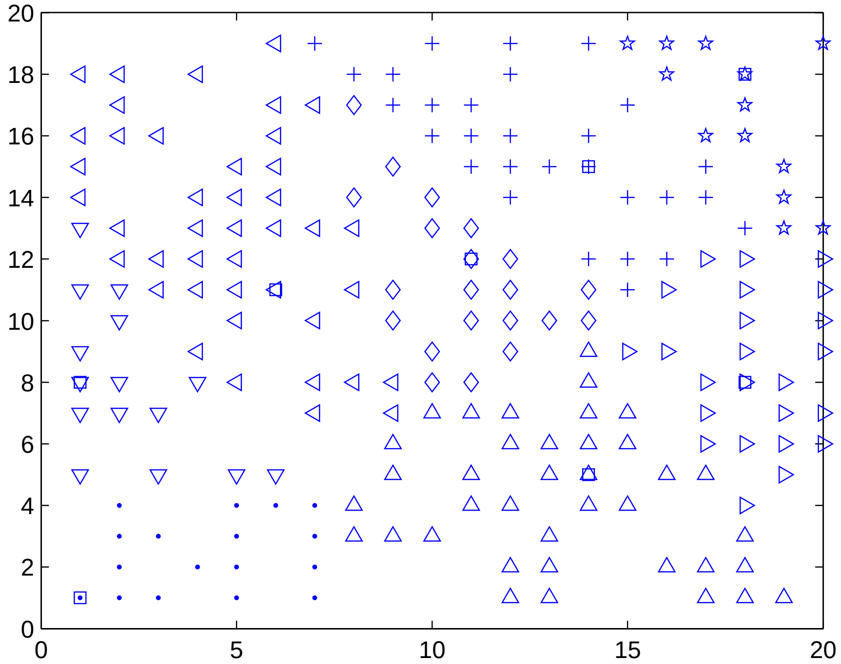

Figure 5.25. Clustering results for round 1500 with algorithm DCE.

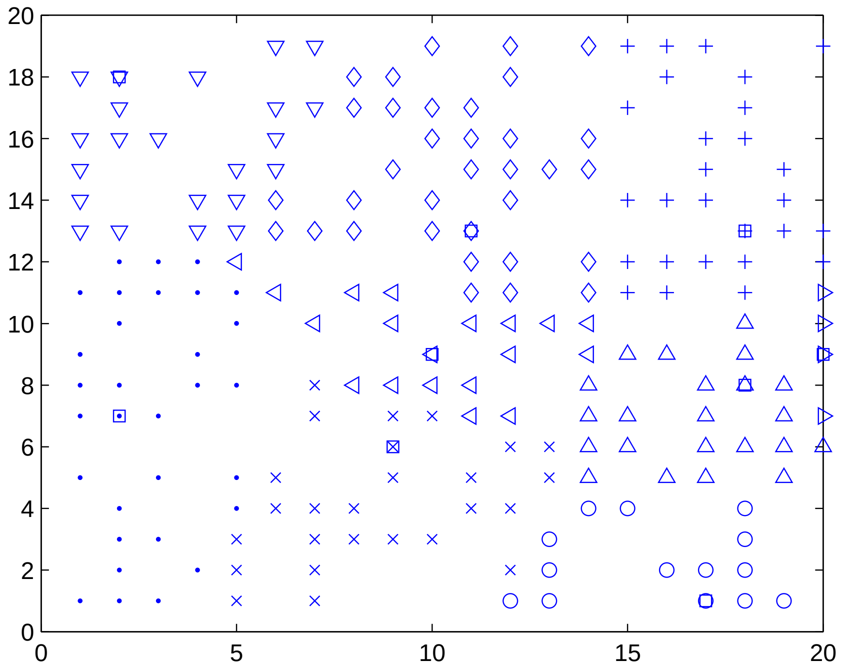

Figure 5.26. Clustering results for round 1800 with LEACH.

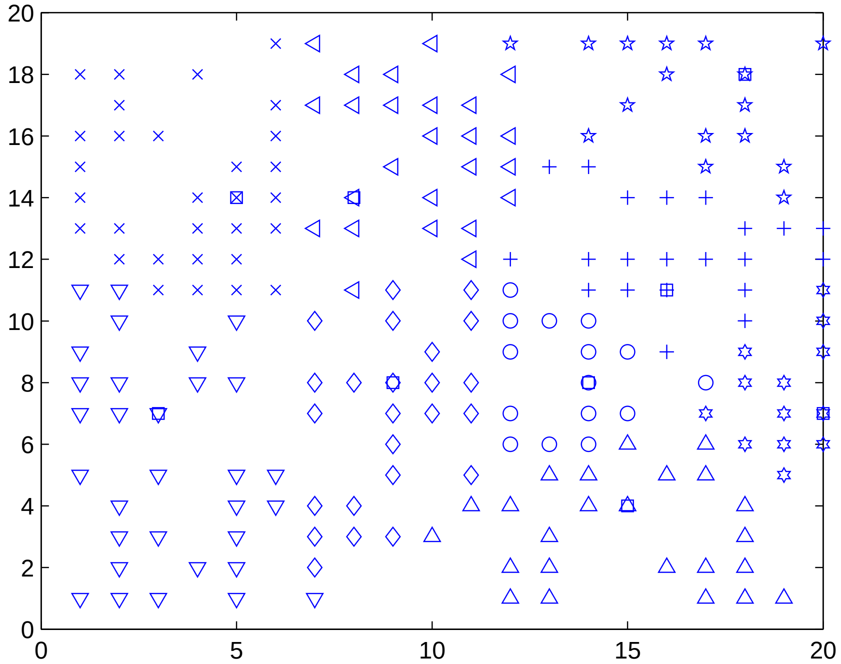

Figure 5.27. Clustering results for round 1800 algorithm DCE.

6. CONCLUSIONS

An enhanced algorithm for wireless network clustering is given in this thesis.The main aim of the protocol is placed on dynamic clustering for adaptive energy consumption.

Wireless sensor networks usually have a random distribution except some special cases in which the positions of sensors are specified before. A random network which contains 200 nodes that are distributed uniformly is used in this thesis . Table 6.1 shows the result for improvement of networks lifetime when it is compared with LEACH. It can be seen from simulation results that final algorithm reaches an improvement of 10 percent for networks lifetime.

Table 6.1. Improvement results.

Number of Rounds	Percentage of improvement	Method
2000	1,5	Algorithm C
2000	6,5	Algorithm E
2000	3	Algorithm D
2000	1	Algorithm R
2000	6,5	Algorithm CE
2000	4	Algorithm DC
2000	10	Algorithm DCE
2000	9,5	Algorithm DCER

REFERENCES

1. Newman M. E. J., “Modularity and community structure in networks”, *The National Academy of Sciences of the USA*, USA, 2006.

2. Chong C.Y., and S.P. Kumar, “Sensor networks: evolution, opportunities,and challenges”, *Proceedings of the IEEE 91*, pp. 1247–1256, 2003.

3. Pottie G.J. , W.J. Kaiser, “Wireless integrated network sensors”, *Communications of the ACM 43*, 2000.

4. Abbasi A. A. , M. Younis, “A survey on clustering algorithms for wireless sensor networks”, *Computer Communications*, Vol. 30, pp. 2826–2841, October, 2007.

5. Heinzelman W. R., and P. Chandrakasan, “ An application-specific protocol architectures for wireless networks”, *IEEE Transactions on Wireless Communications*, pp. 660-670, 2002.

6. Chessa S., and Paolo Santi, “ Fault Diagnosis in Wireless Sensor Networks”, *ERCIM News*, October, 2001.

7. Staddon J., D. Balfanz, and G. Durfee, “ Efficient tracing of failed nodes in sensor networks”, *The First ACM International Workshop on Wireless Sensor Networks and Applications*, Atlanta, USA, September, 2002

8. Meguerdichian S., F. Koushanfar, M. Potkonjak,and M. Srivastava, “ Coverage Problems in Wireless Ad-hoc Sensor Networks”, *Proceedings of the IEEE Infocom*, 2001.

9. Meguerdichian S., F. Koushanfar, G. Qu, and M. Potkonjak, “Exposure In Wireless Ad Hoc Sensor Networks”, *Proceedings of the ACM/IEEE International Conference on Mobile Computing and Networking*, July, 2001.

10. Zhao Y., R. Govindan, and D. Estrin, "Residual energy scans for monitoring wireless sensor networks", *Proceedings of the IEEE Wireless Communications and Networking Conference*, March, 2002.

11. Deb B., S. Bhatangar, and B.I. Nath, "A topology discovery algorithm for sensor networks with applications to network management", *IEEE CAS Workshop on Wireless Communications and Networking*, Pasadena, USA, September, 2002.

12. Zhao J., R. Govindan, " Computing Aggregates for Monitoring Wireless Sensor Networks", *Proceedings of the International Workshop on Sensor Network Protocols and Applications*, April, 2003

13. Ferrari, G., O.K. Tonguz, "Impact of clustering on the BER performance in ad hoc wireless networks", *Communications 2004 International Zurich Seminar*, Zurich, 2004.

14. Akkaya K., M. Younis, "A survey on routing protocols for wireless sensor networks", *Elsevier Journal of Ad Hoc Networks*, 2005.

15. Younis M., M. Youssef, K. Arisha, " Energy-aware management in cluster-based sensor networks", *Computer Networks*, Vol. 43, pp. 649-668, 2003.

16. Dasgupta K., K. Kalpakis, P. Namjoshi, " An efficient clusteringbased heuristic for data gathering and aggregation in sensor networks", *Proceedings of the IEEE Wireless Communications and Networking Conference*, New Orleans, LA, USA, March, 2003.

17. Kwon T., and M. Gerla, "Clustering with power control", *Proc.MILCOM*, Vol. 2, Atlantic City, NJ, November, 1999.

18. Qin Y., J. He, "The impact on throughput of hierarchical routing in ad hoc wireless networks", *Communications, ICC 2005. 2005 IEEE International Conference*, Vol. 5, pp: 3010 -3014, 16-20 May, 2005.

19. Raj R., M.V. Ramesh, and S. Kumar, "Fault Tolerant Clustering Approaches in Wireless Sensor Network for Landslide Area Monitoring", *Proceedings of the 2008 International Conference on Wireless Networks* Vol. 1, pp. 107–113, CSREA Press, July, 2008

20. Beongku A., S. Papavassiliou "An architecture for supporting geomulticast services in mobile ad-hoc wireless networks", *Communications for Network- Centric Operations: Creating the Information Force, IEEE*, Vol. 1, pp. 301 – 305, 28-31 October, 2001

21. Khoshgoftaar T.M., S.V. Nath, S. Zhong, N. Seliya, "Intrusion detection in wireless networks using clustering techniques with expert analysis", *Proceedings Fourth International Conference*, pp. 6, 15-17 December, 2005.

22. Wang Z., L. Liu, M.C. Zhou, N. Ansari, "A Position- Based Clustering Technique for Ad Hoc Intervehicle Communication Systems", *Man, and Cybernetics, Part C: Applications and Reviews IEEE Transactions*, Vol. 38, pp. 201 - 208, March, 2008 .

23. Durresi M., A. Durresi, L. Barolli, F. Hsu, "Parallel Architectures, Algorithms and Networks", *Proceedings, 8th International Symposium*,pp. 6, 7-9 December, 2005.

24. Han S., E. Chan, "Hierarchy Energy Scan in Wireless Sensor Network using in-network aggregation", *the International Journal of Distributed Sensor Networks*, pp. 1- 23, 2009

25. Heinzelman W. R., A. Chandrakasan, and H. Balakrishnan, "Energy-Efficient Communication Protocol for Wireless Microsensor Networks", *Proceedings of the 33rd International Conference on System Sciences*, January, 2000

26. Handy M.J., M. Haase, D. Timmermann, "Low energy adaptive clustering hierarchy with deterministic cluster-head selection", *Proceedings of the 4th International Workshop on Mobile and Wireless Communications Networks*, pp. 368-372, 2002.

27. Hussain S., and A.W. Matin,“ Hierarchical cluster-based routing in wireless sensor networks”, *International Conference on Information Processing in Sensor Networks (IPSN) Work-in-progress track*, Nashville, TN, 19-21 April, 2006

28. Younis, O., and S. Fahmy, “HEED: A Hybrid, Energy-Efficient, Distributed Clustering Approach for Ad-hoc Sensor Networks”, *IEEE Transactions on Mobile Computing* Vol. 4, No. 4, pp. 366–379, Oct-Dec, 2004.

29. Younis, O., and S.Fahmy , “An Experimental Study of Routing and Data Aggregation in Sensor Networks”, *Proceedings of the IEEE International Workshop on Localized Communication and Topology Protocols for Ad Hoc Networks* , Washington, DC, November, 2005.

30. Kroon R. , L.J. Rothkrantz, “Dynamic vehicle routing using an ABC- algorithm”, *Transportation and Telecommunication in the 3rd Millenium*, Prague,Czech Republic, May 26 27, 2003.

31. Tatomir B, H. Dibowsky, L.J. Rothkrantz, “Hierarchical routing in traffic networks”, *16th Belgian-Dutch Conference on Artificial Intelligence*, pp. 75–82, 2004.

32. Qing L., Q. Zhu, and M. Wang, “Design of a distributed energy-efficient clustering algorithm for heterogeneous wireless sensor networks”, *Computer Communication* Vol. 29, pp. 2230-2237, 2006.

33. Smaragdakis G., I. Matta, and A. Bestavros, “SEP: A stable election protocol for clustered heterogeneous wireless sensor networks”, *Proc. of the International Workshop on SANPA*, 2004.

34. Oyman E.I., C. Ersoy, “ Multiple sink network design problem in large scale wireless sensor networks”, *Proceedings of the IEEE International Conference on Communications*, Paris, June 2004.

35. Hou Y.T., Y. Shi, H.D. Sherali,“ On energy provisioning and relay node placement

for wireless sensor networks", *IEEE Transactions on Wireless Communications*, Vol. 4, No. 5, pp. 2579-2590, September 2005.

36. Moy J., "OSPF - Open Shortest Path First", RFC 1583, March 1994.

37. Wang G., G. Cao, T. Porta, and W. Zhang, "Sensor Relocation in Mobile Sensor Networks", *Proceedings of IEEE INFOCOM*, March, 2005.

38. Li X., N. Santoro, and I. Stojmenovic, "Mesh-Based Sensor Relocation for Coverage Maintenance in Mobile Sensor Networks", *Proceedings of the 4th Int. Conf. on Ubiquitous Intelligence and Computing*, pp. 696-708, 2007.

39. Wellenhof B.H., H. Lichtenegger, and J. Collins," Global Positioning System: Theory and Practice Fourth Edition", Springer Verlag, 1997.

40. Kuo Y.C., and S.C. Lin, "Availibility Analysis of Wireless Sensor Networks Based on Sensor Relocation Algorithms", *Journal of the Chinese Institute of Industrial Engineers*, pp. 237–247, 2010.