Survey on Clustering Algorithm for Improved lifetime in HWSN.

Mrs. Impana Appaji
Assistant Professor
Department of Computer Science and Engineering
ATME College of Engineering, Mysore, INDIA

Abstract: clustering algorithm aimed at improving wireless sensor network lifetime in the case of mobile sensor nodes. LEACH-CCH is a modification of the LEACH algorithm, which was developed for stationary networks. An analysis of energy consumption for the LEACH algorithm is presented to identify which data transmissions are most energy expensive for a node throughout its lifetime. LEACH-CCH reduces the energy expended during the costliest data transmission. By predicting the future positions of sensor nodes and restructuring clusters accordingly, an improvement is seen in overall network lifetime when compared with LEACH.

INDEXTERMS: LEACH, DISTRIBUTED CLUSTERING ALGORITHMS, CLUSTER ROUTING PROTOCOL.

I. INTRODUCTION

Wireless sensor networks (WSNs) are collections of low-cost, low-power sensor devices, referred to as nodes, distributed throughout a sensor field. The number of sensor nodes making up a WSN can range from a few hundred to a few thousand. Possible applications include military, agricultural, automotive, space, and environmental concerns making WSNs a topic of large amounts of research over the past few decades. Energy limitations persist due to the interest in minimizing the cost and physical footprint associated with the individual sensor nodes. As a direct consequence of the limited energy characteristics of the sensor nodes that comprise WSNs, the major challenges in this field relate to communication, data storage, and computational capability. Communication methods are necessary to allow data transmission between nodes and to ultimately transmit data collected by sensor nodes to a centralized receiver, known as a base station. A common method of decreasing overall energy consumption is through hierarchical communication routing methods known as clustering.

Clustering protocols are characterized by dividing the network into subsections or ‘clusters’ of fixed or varying numbers of nodes. Data is transmitted from the sensor nodes within a cluster to the base station through an elected clusterhead (CH). Cluster-head responsibilities are usually divided equally among sensor nodes, electing a new cluster-head after each successful data packet transmission, which is referred to as a round. In [1] a pioneering hierarchical cluster protocol known as the Low-Energy Adaptive Clustering Hierarchy (LEACH) is introduced. The LEACH protocol achieved a drastic improvement in overall network lifetime by as much as a factor of eight, when compared to the conventional direct transmission communication and minimumtransmission- energy (MTE) protocols common at the time. Node decay is also seen to occur more uniformly allowing for a more even distribution of sensors throughout the life of the sensor network. Since the development of the LEACH protocol, numerous advances have been made, further improving upon its performance, with a large focus on cluster-head selection algorithms [2]. The majority of related WSN research has been focused on stationary sensor nodes. However, in some applications the sensor nodes will not be stationary. In this case, common protocols such as LEACH, see a significant breakdown in performance. In [3] LEACH-Mobile is introduced as a variation of the LEACH protocol in which considerations are made to better accommodate sensor node movement.

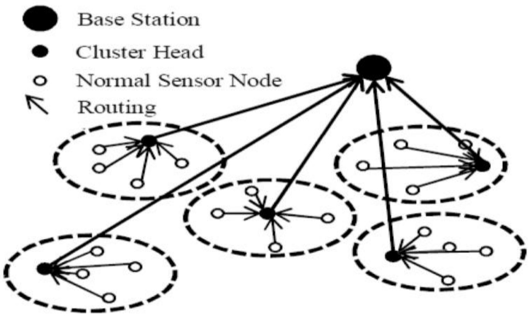

Figure 2: Clustering in Wireless sensor network

One key assumption with LEACH is that any node has the capability to transmit its data directly to the base station. In many applications the sensor nodes will not have this capability. LEACH-Mobile is mainly focused on improving network performance, when this assumption is not valid. In this case a large amount of data is lost when nodes move out of range of their cluster-heads. LEACH-Mobile’s primary focus is on restructuring node clusters once sensor nodes move out of range during LEACH’s steady-state phase. In [3] and [4] it is shown that LEACH-Mobile is an improvement over LEACH when loss of data is considered.

However, it is also shown that energy consumption worsened significantly in LEACH-Mobile due to the increased messaging required to reform node clusters. In [5], the mobility of the cluster-head is considered in the clusterhead election process to further improve upon the performance of LEACH-Mobile.

It is shown that the enhancement improves data transmission by as much as 18% but at the cost of increased computational overhead and energy consumption. This work proposes LEACHCCH (LEACH – Centered Cluster-head), which is a modified LEACH protocol, directed at improving network lifetime when sensor nodes are mobile.

II. REVIEW OF DISTRIBUTED CLUSTERING ALGORITHMS FOR WIRELESS SENSOR NETWORKS

Low Energy Adaptive Clustering Hierarchy (LEACH} is a clustering mechanism which distributes energy consumption all along its network, the network being parted into minor clusters and CHs which are purely distributed in manner and the randomly elected CHs, collect the data from the nodes which are coming under its cluster. The LEACH protocol contains four chief steps for each round: the advertisement phase, the cluster set-up phase, the schedule creation phase and the data transmission phase. During the first step, the advertisement phase, the eligible CH nodes will be delivering an announcement to the nodes coming under them to become a cluster member in its cluster [7]. The nodes will be accepting the offer based on the received signal strength (RSS). In the cluster set-up phase, the sensor nodes will be answering to their selected CHs. In schedule creation step, as the CH accepts response from nodes it have to make a TDMA scheme and send it back to its cluster members to intimate them when they have to pass the data to it. In data transmission step, the data composed by the individual sensors will be given to the CH during their respective time intervals.

The foremost restraint here is that, the radio of the cluster members will be turned off to diminish the energy consumption after the data transmission during particular slot is ended. Here in LEACH clustering protocol, multi-cluster interference problem was solved by using single CDMA codes for each cluster. The energy drain is prohibited for the same sensor nodes which have been elected as the cluster leader using randomization, for each time CH would be altered. The CH is responsible for collecting data from the cluster members and fusing it. Finally, each CH will be forwarding the fused information to the base station. LEACH shows a substantial improvement mainly in terms of energy-efficiency [8-10]. Hybrid Energy-Efficient Distributed Clustering (HEED) is a distributed procedure which selects the CH based on both residual energy and communication cost. Basically, HEED was suggested to avoid the random selection of CHs. Though LEACH protocol is much more energy efficient when compared with its antecedents (discussed below), the primary disadvantage of this method is the random selection of CH. In the worst case, the cluster head nodes may not be consistently distributed among the nodes and it will have its consequence on data gathering. Linked Cluster Algorithm (LCA) is a distributed clustering algorithm that avoids communication collisions among sensor nodes and uses TDMA frames for inter-node communication, with each frame having a time slot for each node in the network for communication. Suggesting cluster formation and CH election algorithms, several papers focuses on single-hop clustering and thereby guarantees that no node will be more than one hop away from leader. In LCA, every nodes necessitates 2n time slots, where n is the number of nodes in the network, to have consciousness of all nodes in its neighborhood [11]-[13]. CLUBS algorithm uses the advantage of local communication to proficiently aggregate the nodes into clusters, in which the time reserved for convergence is proportional to the local density of nodes. In order that the clusters to be advantageous for resource allocation and self-organization, the clustering phenomenon in CLUBS is described by the following: First, every node in the network must apt to some cluster. Second, every cluster should be of equal diameter.

Third, a cluster should have local routing, which means that every node inside the cluster should be able to communicate with each other using only nodes within that same cluster. The CLUBS algorithm forms coinciding clusters, with the maximum cluster diameter of two hops. Every nodes starts competing to form a cluster by choosing random numbers from a fixed integer range [0, R]. Each node counts down from that number silently. If it traces zero without being interrupted, the node becomes a CH and recruits its local neighborhood in to its cluster by broadcasting the recruit message. Energy Efficient Hierarchical Clustering (EEHC) is a distributed and randomized clustering algorithm for WSNs, in which the CHs gather the data about the individual clusters and forward the aggregated report to the base-station. Their method is based on two phases: initial and extended.

The initial phase which is also named as single-level clustering, in which each sensor node proclaims itself as a cluster head with a probability p to the neighboring nodes within its communication range. These CHs are named as volunteer CHs. All the nodes that are within k hops range of a CH receive this announcement either by direct communication or by forwarding. Any node that receives that announcements and is not itself a CH becomes the member of the closest cluster. Forced CHs are sensor nodes that are neither CHs nor fit in to a cluster. If the announcement does not reach to a node within the preset time interval t that is calculated based on duration for a packet to reach a node that is k hops away, the specific node will become a forced CH supposing that it is not within k hops of all volunteer CHs. In the second phase, the technique is prolonged to permit multi-level clustering and commonly builds h levels of cluster hierarchy. Thus, the clustering method is recursively repeated at the level of CHs to form an additional tier. The procedure guarantees h-hop connectivity between CHs and the base-station. Fast Local Clustering Service (FLOC) is a distributed clustering technique that produces non-overlapping clusters and around equal-sized clusters. FLOC achieves locality: effects of cluster formation and faults or changes at any part of the network within almost two units distance. FLOC shows a double-band structure of wireless radio-model for communication. A node can communicate unfailingly with the nodes that are in the inner-band (i-band) range and unreliable communication with the nodes in its outer-band (o-band) range. Hence, the i-band nodes suffer very miniature interference communicating with the CH, thus it is a reliable communication. Messages from o-band nodes are unreliable during data communication and therefore it has the maximum probability of getting vanished during communication. FLOC is fast and scalable, therefore it achieves clustering in O(1) time irrespective of the size of the network. It also displays self-healing capabilities, since the o-band nodes can switch to i-band node in another cluster. It also completes re-clustering within constant time and in a local manner. It also achieve locality, in that each node is only influenced by the nodes within two units.

These structures inspire FLOC algorithm to be suitable for large scale WSNs. Algorithm for Cluster Establishment (ACE) is an extremely uniform cluster formation, self-organizing, slighter overlapping, efficient coverage and emergent cluster forming algorithm for WSNs, which is scale-independent and finishes in time proportional to the deployment density of the sensor nodes irrespective of the overall number of nodes in the network. ACE demands no knowledge of geographic location and necessitates only negligible amount of communication overhead. The important idea of ACE is to assess the potential of a cluster node as a CH before becoming a CH and steps down if it is not the best CH at the moment. The two balanced steps in ACE algorithm is spawning of new clusters and migration of the existing clusters. Spawning is the procedure by which a node becomes a CH. During spawning, when a node approves to become a cluster head, it broadcasts an invitation message to its neighbors.

The neighboring nodes agree such invitation and become a follower of new CH. The principal distinctive feature of ACE is that, a node can be a follower of more than one CH. During migration, best candidate for being CH is selected. Each CH will periodically check all its neighbors to regulate which node is the best candidate to become a cluster head for the cluster. The finest candidate is the node which, if it were to become a cluster head, would have the greatest number of follower nodes with minimum amount of overlap with the prevailing clusters. Once the best cluster head is determined by the current cluster head, it will uphold the best candidate as the new CH and steps down from its CH position [14].

III. THE PROPOSED CC-LEACH CLUSTERING METHODOLOGY

The proposed clustering algorithm is well distributed, where the sensor nodes are positioned randomly to sense the target environment. The nodes are distributed into clusters with each cluster having a CH. The nodes forward the information during their TDMA timeslot to their respective CH which fuses the data to avoid redundant data by the process of data aggregation. The aggregated data is forwarded to the BS. Compared to the existing procedures, the proposed procedure has two distinguishing features. First, the proposed algorithm uses variable transmission power. Nodes nearer to CH use lesser transmission power and nodes far away from CH use extra power for transmission from nodes to CH or vice versa, which can decrease considerable power. Second, CH sends one message for every cluster nodes but many existing algorithms transmits several messages for cluster-setup. The main activity in a WSN is to effectively select a CH. This is attained by using numerous techniques. In the proposed algorithm, CH selection is accomplished with the use of the following parameters (figure 3). A. Highest Coverage In a network of N nodes, each node is assigned an exclusive Node Identity (NID) represented by n, where n=1, 2, 3…., N. The NID merely functions as recognition of the nodes and has no connection with location or clustering. The CH will be located at the center and the nodes will be systematized in to several layers around the CH and these layers are allotted with Layer Number (LN). LN is an integer number beginning from zero. CH gets LN0, nodes adjacent the CH in the next layer are assigned LN1, and so on. In LEACH, the coverage of a sensor node is not taken into account. This is basically significant when a sensor network is used for remote monitoring applications. The nodes with maximum coverage between the cluster nodes are given highest priority to become a CH. Basically HEED was proposed to avoid random selection of CHs. Although LEACH was more energy efficient, the foremost drawback is the arbitrary selection of CH. In HEED, the selection of CH is essentially based on residual energy and communication cost of the nodes. Here the lack of the parameter coverage leads to a main drawback. To overcome these problems, coverage among the nodes is considered to be one of the main parameter in the proposed CC-LEACH algorithm. B. Highest Remaining Energy Remaining energy is defined as to energy remaining within a particular node after some number of rounds. This is normally considered as one of the main parameter for CH selection in the proposed algorithm. LEACH uses much energy for communication among nodes and CHs. It attempts to distribute the loading of CHs to all nodes in the network by switching the cluster heads occasionally. Due to two-hop structure of the network, a node far from CH will have to consume additional energy than a node nearer to CH. This introduces an uneven distribution of energy among the cluster members, disturbing the total system energy and remaining energy. Node death rate is also directly proportional to the remaining energy. It is the measure of the number of nodes die over a time period, from the beginning of the process. When the data rate increases the node death rate also increases. The networks formed by LEACH show periodical variations in the data collection time. This is due to the selection function reliant on the number of data collection process. Since the CH selection of LEACH is a function of the number of completed data collection processes, the number of cluster varies periodically. The same process prevails also in HEED due to enlarged data collection. This increases the node death rate. Therefore, remaining energy is considered as one of the significant parameter for CH selection in the proposed CC-LEACH algorithm. C. Highest Capacity Capacity of a node is the measure of the amount of data processing it can handle compared to other nodes. A node with highest capacity is given priority to become a CH. LEACH uses more energy for communication between nodes and CHs. It tries to distribute the loading of CHs to all nodes in the network by swapping the cluster heads from time to time. The uneven distribution of energy among the cluster members is avoided in HEED as the CH selection is based on residual energy and communication cost. A node with highest residual energy and communication cost becomes a CH, thus the arbitrary selection of CH is avoided. But in the repetition phase, a number of iterations are carried out in order to find the communication cost and selecting a node with better communication cost. This is a peculiar drawback of HEED. In the proposed algorithm, fewer communication energy is necessary. It uses the concept of variable-transmission power in which the transmission power is variable from the lower edge to the higher edge based on the layers. Also in the proposed algorithm, separation among the layers is optimized to use optimum power for each layer. Hence the node with highest capacity is selected as a CH in the proposed CCLEACH algorithm. IV. SIMULATION RESULTS AND DISCUSSIONS For simulation purpose, a sensor network of 50 sensor nodes is randomly organized over a 500 x 500 m2 area. All the sensor nodes are expected to possess equal amount of initial energy. All the simulation mechanisms have been carried out using NS-2. The simulator contains of various components such as deployment component, topology construction component, mobility management component, medium access control component, routing component, energy expenditure computing component and throughput computing component.

IV. LEACH-CCH CLUSTER ROUTING PROTOCOL

A. LEACH Cluster Routing Protocol Overview and Analysis

LEACH forms the basis for LEACH-CCH. LEACH's operation is broken up into rounds, which consist of a set-up phase, in which clusters are formed and a steadystate phase in which data transmission to the base station occurs. At the beginning of the set-up phase, each node determines whether or not to become a cluster-head. This decision is probabilistic and based on the number of rounds since the last time the node was a cluster-head, as well as a predetermined percentage of nodes that should become cluster-heads. Once cluster-heads have been self-nominated, they broadcast an advertisement message to all nodes within the network. Non-clusterhead nodes decide which cluster to join based on the received signal strength of the message. Once this decision is made the non-cluster-head nodes transmit messages to the cluster-head of their choice informing of their decision to join the cluster. Once the cluster is formed, the cluster-head creates and broadcasts a TDMA schedule to the member nodes, which determines when each node will transmit its data during the steady-state phase. The steady-state phase begins with each noncluster head node transmitting its data to the clusterhead at its scheduled time. Once every non-cluster-head node has had a chance to transmit, the cluster-head performs signal processing functions on the data to compress it. The cluster-head then transmits the compressed data to the base station. After this, the round is concluded and the process begins again with a new round. LEACH is well suited for the case in which nodes are stationary, but when nodes are moving the performance of LEACH is significantly reduced in terms of network lifetime.

Recognizing that LEACH was not intended for such operation, we were interested in the possibility that a few slight modifications could increase the performance to more acceptable levels. LEACH was simulated, as described in section III.B, in part to observe which data transmissions were most costly when sensor nodes were mobile. Simulations indicated throughout the network lifetime, 72% of a node's available energy is spent transmitting while 28% is consumed receiving transmissions. The most expensive transmission occurs during the steady-state phase, when a node is transmitting data to its cluster-head, as shown in Table I. This transmission accounts for 69% of a single node's available energy throughout the network lifetime. This is a significant increase when compared to 57% in the case of stationary sensor nodes. One might anticipate that the majority of a node's energy would be consumed when transmitting to the base station during its rounds as cluster-head; however, a node is only going to be cluster-head for approximately 5% of the rounds during its lifetime, but it will be transmitting data to its cluster-head for the remaining 95%.

B. LEACH-CCH Cluster Routing Protocol

The motivation for LEACH-CCH is to improve the network lifetime of LEACH in the case of mobile sensor nodes. A flowchart providing a high-level view of a LEACH-CCH round is given in Fig. 1. When nodes are moving, the major cause of degraded performance is because the clusters formed during the set-up phase will break apart throughout the steady-state phase of a round as nodes move away from one another, as illustrated in Fig. 2, Fig. 3, and Fig. 4. This breakdown of clusters increases the energy expended by the noncluster- head nodes by increasing the distance each node must transmit to reach its elected cluster-head. LEACH-CCH makes a modification to the LEACH protocol during the set-up phase of a round aimed at reducing the effect of this cluster decay. During the setup phase of LEACH, cluster-heads are self-elected based on a probabilistic calculation, which is a function of the number of rounds since the node was a clusterhead as well as a predetermined percentage of nodes that should become cluster-heads, as detailed in [1]. Each cluster-head, then transmits an advertisement message to all nodes within the network. Non-clusterhead nodes receive this message and select the nearest cluster-head based on the signal strength of the advertisement message. Each non-cluster-head node then transmits a cluster-head selection message indicating its decision to join the cluster. LEACH-CCH assumes this selection message also contains information including the non-cluster-head node's current velocity and position.

At this point, clusters have been formed as shown in the network depicted in Fig. 5. In LEACH-CCH, a restructuring of the clusters is conducted at this point, as detailed in Algorithm 1, in which a new cluster-head is selected based on which node is nearest to the center of the cluster, as shown in Fig. 6. This reduces the distance between the majority of the nodes in the cluster and the cluster-head, which reduces the energy expended during data transmission to the cluster-head. As nodes move during the steadystate phase, the center of the cluster will also move. This makes it beneficial to predict the location of nodes throughout this phase of the round. Assuming the position and velocity of nodes are known by the clusterhead during the formation of clusters and also assuming velocity is constant. The cluster-head can determine where each node will be at different times in the steady-Fig. 2. Mobile sensor node clusters at the beginning of the LEACH steady-state phase at time *t*. (Larger icons indicate cluster-heads) Fig. 4. Mobile sensor node clusters during the LEACH steady-state
phase at time *t + 10*.

for timesteps *i* to *tss*
calculate cluster center of mass, *R*
for each node in cluster, *ni*

ni movement advanced *i* time steps
centerDistances_*ni* += *ni* distance to *R*
end loop
end loop
let bestDistance = current CH's distance to *R*
let bestNode = current CH
for each node in cluster, *ni*
avgCenterDistance = centerDistances_*ni*/*tss*
if avgCenterDistance < bestDistance
bestNode = *ni*
bestDistance = avgCenterDistance
end if
end loop
New cluster-head ≡ bestNode

state phase. Before transmitting the TDMA schedule, the cluster-head determines the number of time steps required for all cluster nodes to transmit their data to the cluster-head, *tss*, and steps forward in time to determine the position of each node throughout the steady-state phase. At each time step, the cluster-head calculates the center of mass of the cluster and records each node's distance from the center until the final time step is reached. The mean of these distances is taken, and the node with the shortest average distance becomes the new cluster-head. LEACH-CCH makes no consideration for the number of times a node has been selected as the replacement cluster-head, meaning a node could potentially become the replacement clusterhead multiple rounds in a row. When simulating LEACH-CCH, modifications were made to limit the number of rounds a node could become replacement cluster-head, but it was observed that the network lifetime was best when this consideration what not made. Often times the original cluster-head is the best candidate, and in this situation the original node remains the cluster-head. In the event of a tie for replacement cluster-head, a random selection is made between the two.

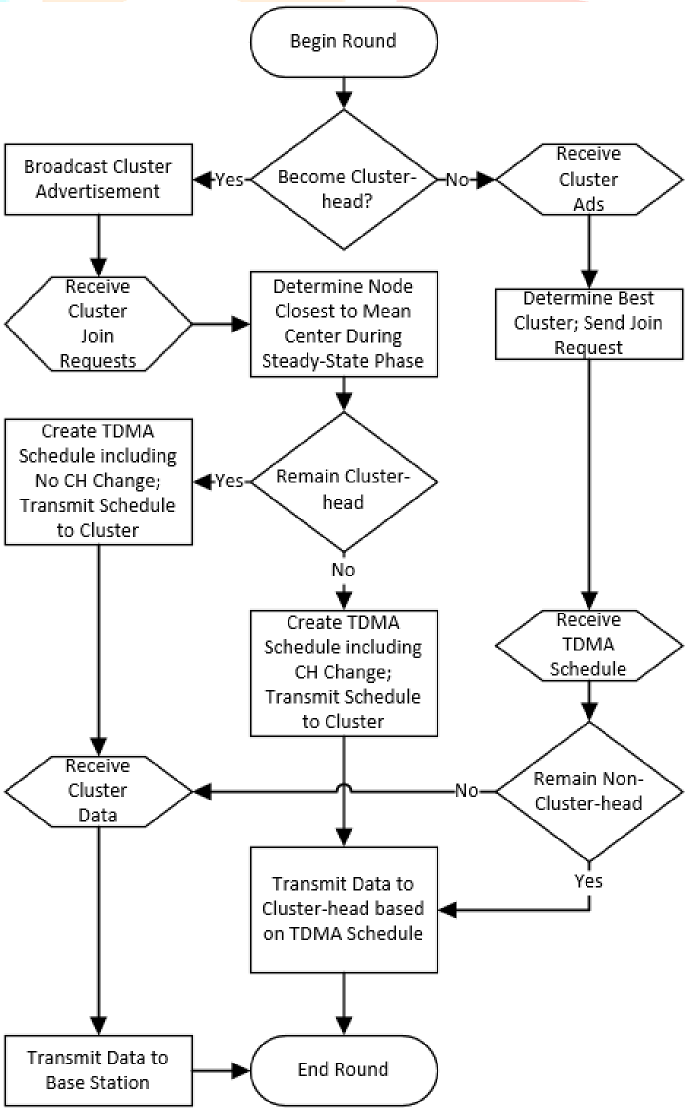

Fig.2. Flowchart detailing a round of the LEACH-CCH algorithm

V. CONCLUSION AND FUTURE WORK

It has been demonstrated that LEACH-CCH achieves improved network lifetime over LEACH in both cases of stationary and mobile sensor nodes. LEACH-CCH achieves this improvement by restructuring sensor node clusters before the beginning of the steady-state phase. By reselecting cluster-heads based on which node is closest on average to the center of the cluster, the distances non-cluster-head nodes must transmit their data are improved. Future work will consider forming clusters based on predicting future node movement to minimize cluster decay throughout the steady-state phase of the round. Future work will also observe the scenario in which cluster decay becomes more of an issue when radio transmission capability is limited, in which case a node can travel out of range of its clusterhead resulting in a loss of data.

REFERENCES

[1] W. R. Heinzelman, A. Chandrakasan and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *System Sciences, 2000. Proceedings of the 33rd Annual Hawaii International Conference on*, 2000.

[2] L. M. C. Arboleda and N. Nasser, "Comparison of Clustering Algorithms and Protocols for Wireless Sensor Networks," *2006 Canadian Conference on Electrical and Computer Engineering*, Ottawa, Ont., 2006, pp. 1787-1792.

[3] Y. Chung and D. Kim, "Self-Organization Routing Protocol Supporting Mobile Nodes for Wireless Sensor Network," in *First International Multi-Symposiums on Computer and Computational Sciences*, 2006.