An Improved Cluster Head Election Technique in Wireless Sensor Networks

Akhil Garg¹, Tarun Kumar²

¹M.tech Scholar, Department of Comp Sci. & Engg., KUK
²Asstt. Prof., Department of Comp Sci. & Engg., KUK

Abstract: A wireless sensor network is an autonomous system of sensor nodes. It has a Base Station and sensor nodes. Sensor nodes collect data from their environment and send it to the Base Station. Heterogeneous sensor network contains high energy sensor nodes as well as low energy nodes. A single-tier network can cause the gateway to overload with the increase in sensors density. Such overload might cause latency in communication and inadequate tracking of events. In addition, the single-tier architecture is not scalable for a larger set of sensors covering a wider area of interest because the sensors are typically not capable of long-haul communication. Hierarchical clustering is particularly useful for applications that require scalability to hundreds or thousands of nodes. Scalability in this context implies the need for load balancing and efficient resource utilization.

All nodes in a network can be organized in hierarchical structures called clusters. Each cluster consists of a cluster head and several member nodes. The election of cluster head by normal nodes is a tedious task. Using Fuzzy Logic this task can be done easily. So Fuzzy is used for cluster head election. The member nodes collect data and send it to their cluster heads. The cluster head aggregates and transmits the data to the Base Station. The energy consumption of cluster heads is higher than that for member nodes. Clustering algorithms are required which can efficiently utilize the energy of nodes so that life of network can be increased.

Keywords: WSNs, Clustering, SN, CH, Fuzzy Logic

I. INTRODUCTION

WSNs consist of a large number of limited capabilities (power and processing) Micro Electro Mechanical Systems (MEMS) capable of measuring and reporting physical variables related to their environment. A WSN consists of spatially distributed autonomous sensors to cooperatively monitor physical or environmental conditions, such as temperature, sound, vibration, pressure, motion or pollutants. Sensor networks are being deployed for a wide variety of applications [7], including by military applications such as battlefield surveillance and is now used in many industrial and civilian application areas, environment and habitat monitoring, healthcare applications, home automation, and traffic control. In surveillance applications, sensors are deployed in a certain field to detect and report events like presence, movement, or intrusion in the monitored area. Data collected by sensors are transmitted to a special node equipped with higher energy and processing capabilities called “Processing Node” (PN) or “sink”. The PN collects, filters, and compiles data sent by sensors in order to extract useful information.

A. Sensor Network Architecture

We consider the sensor network architecture depicted in Figure 1.1. In the architecture SNs are grouped into clusters controlled by a single command node. Sensors are only capable of radio-based short-haul communication and are responsible for probing the environment to detect a target/event. Every cluster has a gateway node that manages sensors in the cluster. Clusters can be formed based on many criteria such as communication range, number and type of sensors and geographical location. Sensors receive commands from and send readings to its gateway node, which processes these readings.

Fig. 1 Sensor Network Architecture
Gateways can track events or targets using readings from sensors in any clusters as deemed by the command node. However, sensors that belong to a particular cluster are only accessible via the gateway of that cluster. Therefore, a gateway should be able to route sensor data to other gateways. Gateway nodes interface the command node with the sensor network via long haul communication links.

The gateway node sends to the command node reports generated through fusion of sensor readings, e.g. tracks of detected targets. The command node presents these reports to the user and performs system-level fusion of the collected reports for overall situation awareness.

II. CLUSTERING IN WSNs

The major advantage of WSN is the ability to deploy it in an ad-hoc manner [4], as organizing these nodes into groups pre-deployment is not feasible. For this reason, a lot of research has been conducted into ways of creating these organizational structures (or clusters) [5]. A clustering scheme divides the sensor nodes in a WSN into different virtual groups, according to some set of rules. In a cluster structure, sensor nodes may be assigned a different status or function, such as cluster head or cluster member [21]. We can see in the Figure 1.2, the architecture of a generic WSN, and examine how clustering is an essential part of the organizational structure [5].

Sensor Nodes: Sensor nodes are the building blocks of a WSN. They can play multiple roles in a WSN, such as simple sensing, data processing, data storage and routing. Clusters: Clusters are the organizational unit of WSNs. The dense nature of WSNs requires them to be broken down into clusters to simplify tasks such as routing. Cluster heads: Cluster head is the organizational leader of a cluster. It organizes the activities in a cluster.

The activities include data-aggregation, diffusion, organizing the communication schedule of the cluster, etc. Base Station: The base station is often located far from the network. It provides the communication link between the WSN and the end-user. End User: The data obtained from sensor network can be used for a wide-range of applications. A particular application can make use of the network data over the internet, using a PDA, or even a personal computer. In a queried sensor network, queries are generated by the end user.

A. Clustering Algorithms

Many algorithms have been proposed for routing in WSN. Clustering algorithms have gained popularity in this field. Clustering algorithms can be classified as:

1) Distributed algorithm,
2) Centralized algorithm,
3) Hybrid algorithm

In distributed clustering techniques, any node can choose itself as a CH or join an already formed cluster on its own initiative, independent of other nodes. Distributed clustering techniques are further classified into four sub types based on the cluster formation criteria and parameters used for CH election as identity based, neighborhood information based, probabilistic and iterative. In centralized methods [22], the BS requires global information of the network to control the network. CHs are elected by the base station. Hybrid schemes are composed of centralized and distributed approaches. In a hybrid environment, distributed approaches...
are used for coordination between CHs, and centralized schemes are followed for CHs to build individual clusters. In design of routing protocols for WSN, clustering algorithms have following advantages:

1) Clustering reduces number of nodes taking part in long distance transmission.
2) Clustering algorithms are scalable for large number of nodes.
3) They reduce communication overhead.
4) Energy is utilized properly by the use of clustering algorithms.

Il. Literature Review

A. Cluster-Construction Based Clustering Routing Protocols

1) Leach: LEACH (low-energy adaptive clustering hierarchy) proposed by Heinzelman et al. [1], is one of the pioneering clustering routing approaches for WSNs. It divides the protocol operation into rounds, and each round is subdivided into two phases: setup and steady-state phase. In the setup phase, the nodes create clusters and elect CH. The node becomes a CH for the current round if the number is less than the following threshold:

\[ \text{CHprob} = \frac{E_{\text{residual}}}{E_{\text{max}}} \]

where \(E_{\text{residual}}\) is the estimated current energy of the node, and \(E_{\text{max}}\) is a reference maximum energy, which is typically identical for all nodes in the network. The value of CHprob, however, is not allowed to fall below a certain threshold that is selected to be inversely proportional to \(E_{\text{max}}\). Afterwards, each node goes through several iterations until it finds the CH.

2) Heed: Hybrid Energy-Efficient Distributed clustering (HEED) [2], introduced by Younis and Fahmy, is a multi-hop WSN clustering algorithm which brings an energy-efficient clustering routing with unambiguous thought of energy. Different from LEACH in the manner of CH election, HEED does not select nodes as CHs arbitrarily. In HEED, CHs are periodically elected based on two important parameters: residual energy and intra-cluster communication cost of the candidate nodes. Initially, in HEED, a percentage of CHs among all nodes, \(C_{\text{prob}}\), is set to assume that an optimal percentage cannot be computed a priori. The probability that a node becomes a CH is:

\[ \text{CHprob} = \frac{E_{\text{residual}}}{E_{\text{max}}} \]

where \(E_{\text{residual}}\) is the estimated current energy of the node, and \(E_{\text{max}}\) is a reference maximum energy, which is typically identical for all nodes in the network. The value of CHprob, however, is not allowed to fall below a certain threshold that is selected to be inversely proportional to \(E_{\text{max}}\). Afterwards, each node goes through several iterations until it finds the CH.

3) Ti-Leach: Two-Level Hierarchy LEACH (TL-LEACH), presented by Loscri et al. [8], is an addition to the algorithm of LEACH. TL-LEACH uses the following two techniques to achieve energy and latency efficiency: randomized, adaptive, self-configuring cluster formation and localized control for data transfers. A CH collects data from MNs as original LEACH, but instead of transmitting data to the BS directly, it uses a part of CHs that lies between the CH and the BS as a relay station. The algorithm has four basic phases: advertisement phase, cluster setup phase, schedule creation and data transmission. In the first phase, each node decides whether it become a primary CH, secondary CH or ON in each round which is the same as that of LEACH. If a node is elected a primary CH, it must advertise other nodes. The mechanism used in this phase is carrier sense multiple access (CSMA). Thereafter, secondary CH nodes send the advertisement to the ONs. In this phase, each secondary CH decides to which primary CH it belongs and sends an advertisement message to its primary CH. Each ON must decide which secondary CH it belongs to and informs it through an opposite message. In the third phase, each primary CH creates a TDMA schedule assigning each node in its group a slot to transmit.

4) Eecs: Energy Efficient Clustering Scheme (EECS), proposed by Ye et al. [3,10], is a clustering algorithm which better suits the periodical data gathering applications. EECS is a LEACH-like scheme, where the network is partitioned into several clusters and single-hop communication between the CH and the BS is performed. In EECS, CH candidates compete for the ability to elevate to CH for a given round. This competition involves candidates broadcasting their residual energy to neighboring candidates. If a given node does not find a node with more residual energy, it becomes a CH. Different from LEACH for cluster formation, EECS extends LEACH by dynamic sizing of clusters based on cluster distance from the BS.

5) Eeuc: Energy-Efficient Uneven Clustering (EEUC) algorithm, proposed by Li et al. [18], is a distributed competitive and clustering algorithm, where CHs are elected by confined competition, which is not like LEACH. Each node has a pre-assigned viable range, which is smaller as it gets close to the BS. This gives EEUC an uneven clustering approach for the determination of balancing energy consumption among CHs. In the process of CH election, each node generates a random number, and only the node whose number is greater than a threshold will be activated for CH election by broadcasting compete message within a competition radius which is determined by its distance to the BS.

6) Ace: Algorithm for Cluster Establishment (ACE) [20], given by Chan and Perrig, employs an emergent algorithm, which is any calculation that achieves formally or stochastically probable global effects, by communicating directly with only a bounded number of instant neighbors and without the use of central control or global visibility. One of the main individual characteristics of growing protocols over other localized protocols is the survival of feedback during protocol operation.
7) **Bcdcp**: Base-Station Controlled Dynamic Clustering Protocol (BCDCP), introduced by Muruganathan et al. [16], is a centralized clustering routing protocol with the BS capable of difficult computation. The main idea of BCDCP is the cluster formation where each CH serves an almost equal number of MNs to balance CH overload and uniform CH placement throughout the network.

### B. Data-Transmission Based Clustering Routing Protocols

1) **Pegasis**: Power-Efficient Gathering in Sensor Information Systems (PEGASIS), proposed by Lindsey et al. [12], is an improvement of LEACH. The main idea of PEGASIS is for each node to only communicate with their close neighbors and take turns being the leader for transmission to the sink. In PEGASIS, the locations of nodes are random, and each sensor node has the ability of data detection, wireless communication, data fusion and positioning. Energy load is distributed evenly among the sensor nodes in the network.

2) **Teen**: Threshold sensitive Energy Efficient sensor Network protocol (TEEN) [15], proposed by Anjeshwar and Agrawal, is a hierarchical protocol whose main goal is to cope with sudden changes in the sensed attributes such as temperature. The protocol combines the hierarchical technique in line with a data-centric approach. The nodes sense their environment continuously, but the energy consumption in this algorithm can potentially be much less than that in the proactive network, because data transmission is done less frequently. In TEEN, a 2-tier clustering topology is built and two thresholds, hard threshold and soft threshold, are defined. The former threshold is a threshold value for the sensed attribute. It is the absolute value of the attribute beyond which, the node sensing this value must switch on its transmitter and report to its CH. The latter threshold is a small change in the value of the sensed attribute which triggers the node to switch on its transmitter and transmit.

3) **Apteen**: The Adaptive Threshold sensitive Energy Efficient sensor Network protocol (APTEEN) [14], introduced by Manjeshwar and Agrawal, is an extension to TEEN and aims at both transmitting periodic data and reacting to time critical events. APTEEN, on the other hand, is a hybrid protocol that changes the periodicity or threshold values used in TEEN according to the requirement of users and the type of the application. APTEEN is based on a query system which allows three types of queries: historical, on-time, and persistent which can be used in a hybrid network. Moreover, QoS requirements are introduced for the on-time queries and minimum delay is achieved by a TDMA schedule with a special time slot assignment manner.

4) **Ttdd**: The Two-Tier Data Dissemination (TTDD) approach, presented by Luo et al. [13], is a low-power protocol for efficient data delivery from multiple sources to multiple mobile sinks. It exploits a geographic routing based on grid of cells as the routing method. Instead of passively waiting for queries from sinks, sensor nodes can proactively establish a structure to set up forwarding information. Ultimately, the sensing field is figured as a set of grid points.

5) **Ccs**: The Concentric Clustering Scheme (CCS) has been proposed in [9] by Jung et al. to reduce the energy consumption loopholes in PEGASIS. The main idea of CCS is to consider the location of the BS to enhance its performance and to prolong the lifetime of the network.

6) **Hgmr**: Hierarchical Geographic Multicast Routing (HGMR), proposed in [23] by Koutsonikolas et al., is a location-based multicast protocol. This protocol seamlessly incorporates the key design concepts of the Geographic Multicast Routing (GMR) [8] and Hierarchical Rendezvous Point Multicast (HRPM) protocols and optimizes them by providing forwarding energy efficiency as well as scalability to large-scale WSNs. HGMR starts with a hierarchical decomposition of a multicast group into subgroup of manageable size by means of the key concept of mobile geographic hashing of HRPM. Within each subgroup, HGMR adopts the local multicast scheme of GMR to forward data packets along multiple branches of the multicast tree in one transmission. In HGMR, the multicast group is divided into subgroups using the mobile geographic hashing idea: the deployment area is recursively partitioned into a number of $d^2$ equal-sized square sub-domains called cells, where $d$ is decomposition index depending on the encoding overhead constraints, and each cell comprises a manageable-sized subgroup of members. In each cell there is an Access Point (AP) responsible for all members in that cell, and all APs are managed by a Rendezvous Point (RP). In [37] a new approach for WSN clustering using Fuzzy logic is studied. The parameters which are used in the algorithm are residual energy, distance from the base station and distance from CH. The goal of this paper was to study a new logic to select the best cluster heads that combine different criteria modifying the energy efficiency of cluster heads. Also, using the proposed algorithm, the energy consumption per round is evaluated and compared with some previous methods which have proved better performance and improves energy consumption. The scheme can be effective to use in large scale networks. There is always a scope of methods for more energy efficient networking in WSN clustering.
IV. PROPOSED RESEARCH WORK

A. Problem Definition
A wireless sensor network is an autonomous system of sensor nodes. It has a Base Station and sensor nodes. Sensor nodes collect data from their environment and send it to the Base Station. Heterogeneous sensor network contains high energy sensor nodes as well as low energy nodes. A single-tier network can cause the gateway to overload with the increase in sensors density. Such overload might cause latency in communication and inadequate tracking of events. In addition, the single-tier architecture is not scalable for a larger set of sensors covering a wider area of interest because the sensors are typically not capable of long-haul communication. Hierarchical clustering is particularly useful for applications that require scalability to hundreds or thousands of nodes. Scalability in this context implies the need for load balancing and efficient resource utilization. All nodes in a network can be organized in hierarchical structures called clusters.

B. Objectives
In this research we study few of clustering Routing techniques in WSN. The summarized objective of work for the dissertation as follows.
1) To find an effective clustering method in WSN in terms of energy efficiency.
2) To enhance the life time of network.
3) Better election of cluster head (CH) to save.
4) To compare the existing technique with the proposed technique using the different parameters.

C. Improved Cluster-based approach for energy-efficiency in the WSNs
1) CH election: After adjusting the transmission power, each node generates a random number (μ), which ranges from 0 to 1. Then, the node decides to become a CH by comparing μ with the $T(n)$, which is computed by means of Equation 1. If $\mu$ is less than $T(n)$, the node becomes a CH for the current round.

$$T(n) = \eta \left( \frac{\sigma^2}{\alpha} \right) + \mu \left( 1 - e^{\frac{-\sigma^2}{\alpha}} \right)$$

Where $\eta$ and $\alpha$ are weights to give importance, the sum is exactly 1. The Residual Energy is denoted as $RE$, and $\sigma$, means the energy variance, which is used to produce better CH candidates.
Equation 1 uses a gauss function, due to the fact that has better result in terms of energy efficiency and representation in the context of an imprecise environment.
for a join message from the non-CH nodes. However, if the CHs do not receive a join message, this CH should not become CH. Algorithm 1 describes the steps for CH election and cluster formation.

2) Cluster Formation: Now in present work author had used three linguistic input variables of the system are residual energy, distance from non-CH to base station (expressed in meters) and distance from CH to base station(expressed in meters). The specifications related for the input and output functions of the system and their respective Linguistic Values (LV) are as follows:
Residual energy: $u=[0,100]: LV = low, average, high$;
Distance: $u=[0,100]: LV = small, average, big$; 
Dist_base: $u=[0,100]: LV = small, average, big$;
Probability: $u=(0,1]: LV = very high, Medium high, high, Moderately medium, fairly medium, medium, Moderately low, low, very low$.For the representation of the linguistic states (low, high, small and large) of the input variables, the degrees of membership to these sets must remain constant for certain values of the universe of discourse.

3) The Pseudo code of Proposed Model is as Follows:
Step1: Start
Step 2: Create a Network
Step 3: Create Clusters from network using:
A CH is selected from the SNs.
Based on last step, Non-CHs select the best CH by considering a multiple metrics i.e. residual energy, distance from non-CH to base station and distance from CH to base station using the concept of Fuzzy logic and Cluster is created.
Step 4: Stop
The Flow Chart of the proposed model is given in figure 3.

Fig. 3 Flow Chart-based approach for energy-efficiency
V. RESULTS

A. Performance Evaluation

The basic parameters used for simulations are listed in table 1.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Field Size</td>
<td>100m X 100m</td>
</tr>
<tr>
<td>Location of Base Station</td>
<td>Random</td>
</tr>
<tr>
<td>No. of Nodes</td>
<td>50</td>
</tr>
<tr>
<td>Probability of cluster</td>
<td>0.1</td>
</tr>
<tr>
<td>Initial Energy of sensor node</td>
<td>20 J</td>
</tr>
<tr>
<td>The Data packet Size</td>
<td>4000 bits</td>
</tr>
<tr>
<td>Delta T</td>
<td>10</td>
</tr>
<tr>
<td>( \theta )</td>
<td>0.4</td>
</tr>
<tr>
<td>( \alpha )</td>
<td>0.6</td>
</tr>
<tr>
<td>( E_{fs} )</td>
<td>10 J/bit/m²</td>
</tr>
<tr>
<td>( E_{mp} )</td>
<td>0.0013 J/bit/m²</td>
</tr>
</tbody>
</table>

Based on these parameters author will carry out the simulations. These parameters are taken after studying different research papers used in Wireless sensor network.

Figure 4 shows the comparison of energy consumed by existing and new scheme. It shows that the new scheme is more energy efficient than the traditional schemes.

![Fig. 4 Comparison of Energy Consumption](image)

Figure 5 shows the comparison of first dead node in existing and new scheme. It shows that the first dead node in new scheme is at round no. 380 where as in existing scheme it is at round no. 355.

![Fig. 5 Comparison of First Dead node](image)
Figure 6 shows the comparison of Total no. of dead nodes in existing and new scheme. It shows that the dead nodes in new scheme are 25 where as in existing scheme these are 32 in nos.

Finally table 2 summarizes the comparison between two schemes based on various parameters. Both the schemes had been run many times and finally the comparison table is drawn.

<table>
<thead>
<tr>
<th>Considerations</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy Comparison</td>
<td>New Scheme is % better than existing scheme.</td>
</tr>
<tr>
<td>First Dead</td>
<td>First dead nodes in new scheme is far late than existing scheme.</td>
</tr>
<tr>
<td>Total No. of Dead Node</td>
<td>Total dead nodes in new scheme 45% (approx.) less than existing scheme.</td>
</tr>
</tbody>
</table>

VI. CONCLUSIONS

A. Conclusion

For a better cluster formation the concept of fuzzy logic is used in which non-CHs select the best CH by considering a multiple metrics, i.e. residual energy, distance from CH to BS and a distance from non-CH to BS. This is done because in existing authors were not considering the distance between non-CH and BS. Then, non-CHs compute a probability value to each CH candidate. The non-CH chooses the CH with a higher probability value and sends a join message to CH.

The use of fuzzy logic is suitable, whenever it is not possible to use a mathematical model for the system. Additionally, fuzzy can reduce the complexity of the model, computational effort and memory. Energy consumption is affected by message communication between nodes, so this technique is efficient than traditional protocol.

REFERENCES


