A Comparative Study of Existing Cluster-Based Routing Protocols in Wireless Sensor Networks

Pooja Mishra\textsuperscript{1} and Sanjiv Sharma\textsuperscript{2}

\textsuperscript{1}M. Tech. (IT) Scholar, Department of CSE/IT, MITS, Gwalior, India
\textsuperscript{2}Assistant Professor, Department of CSE/IT, MITS, Gwalior, India

\textsuperscript{1}mishrapooja427@gmail.com
\textsuperscript{2}er.sanjiv@gmail.com

Abstract

Wireless Sensor Network (WSN) has grown vastly in the past few years and pointing out the crucial need for scalable and energy efficient network. Routing in WSN is a challenging task because of limited resources of available energy and processing power. On the basis of network structure, routing protocol in WSN can be classified into two different categories such as flat routing and cluster-based or hierarchical routing. In literature review, researchers proposed a number of routing protocols in WSN. In which cluster-based routing protocols shows better performance as compare to other routing protocols. Therefore, the objective of this paper is to provide a review on some efficient cluster-based routing protocols with advantages and limitations.

Keywords: Wireless Sensor Network (WSN), Clustering, Network lifetime, Energy efficiency

1. Introduction

Wireless Sensor Networks (WSNs) [3] is a densely deployed communication network, which consists of several nodes and responsible for gathering information from their surroundings. All of this information is sent to the sink node so that it can be accessed by the end user only. WSN has several applications such as military reconnaissance, industrial automation, security surveillance, disaster management, habitat monitoring, medical and health care monitoring, environmental/earth sensing [2,4] etc.

A WSN node combines sensing, computation and communication into a single tiny device. Prolonged network lifetime, scalability, load balancing and minimum energy consumption are important need for various WSN applications. Existing work reveal that clustering is a standard technique for achieving efficient and scalable performance in this type of networks [3]. The formal method of direct transmission in which nodes quickly die down because each node transmits data directly to base station (BS) which is located at a far distance. So the elegant approach is clustering in which network is divided into disjoint groups and each group is managed by a representative node known as cluster head (CH) and others are member nodes.

In two levels hierarchy, only CH nodes communicate directly to the BS. This two level hierarchical approach is defined by the most basic protocol known as LEACH (Low Energy Adaptive Clustering Hierarchy) by W. R. Heinzelman et al. [11]. It was the first cluster-based distributed approach in WSN.

The rest of paper organized as follows: The section 2 presents related work of clustering routing protocols. Section 3 presents clustering objectives and advantages, some important characteristics of cluster formation. Section 4 presents classification of different clustering protocols with some clustering issues and
challenges. Section 5 presents comparison of different cluster-based routing protocols. Finally, section 6 presents conclusion that cluster-based routing protocol is used to improve the performance of WSNs.

2. Related Work

In the past, large amount of research carried through cluster based algorithms in WSNs. A survey on energy efficient clustering routing protocol in WSN by V. kumar et al. [5] reported LEACH and its descendants with different energy efficient clustering algorithms for improving the network lifetime of WSNs.

S. R. Prabhu et al. [6] Authors presented some frequently used distributed clustering algorithm and shows the difference among them based on some metrics such as cluster count, stability, cluster head mobility, cluster-head role, clustering objective, cluster-head selection, merits and limitations.

S. K. Gupta et al. [7] authors presented energy efficient algorithms based on cluster head selection techniques, these are: probability based and non-probability based. In which, probability based technique are less energy consumable. In this paper, several clustering approaches are elaborated with comparison based on clustering properties and various parameters such as cluster head mobility and location awareness.

V. Kumar et al. [8] authors presented different hierarchical clustering algorithms. These are most focused on data aggregation/fusion so that amount of data transmitted to the base station is reduced and improve the energy efficiency of WSNs. Furthermore, compared the clustering algorithms based on cluster count, heterogeneity, cluster overlapping etc. exists in literature of WSNs.

Xuxun Liu [9] presented some classical WSN clustering routing protocols in detail. Furthermore, compared all routing protocols based on some primary metrics and examines how the several challenges deal with the design of effective, robust, and scalable routing protocols.

G. Sikander et al. [10] presented various cluster-based routing protocols in WSN and compared different clustering methods on the basis of various performance measures such as cluster stability, scalability, delivery delay, complexity, energy efficiency and load balancing.

3. Clustering

In clustering, the sensor network is divided into different clusters. Each cluster has a representative node known as CH and other are cluster members. Member nodes do not communicate directly with the sink node or BS. They have to forward the aggregated data to the CH. The CH will perform aggregation of the received data from member nodes and sends it to the BS.

Assume that N nodes are spread in a field. Our goal is to identify a set of CHs which cover the entire network efficiently. Each node is associated with exactly one cluster. The node can directly communicate with its CH.

The following requirements must be met [1]:
1) Clustering is completely distributed.
2) Clustering should be terminated within a fixed number of iterations. At the end of each node is either a CH or a cluster member.
3) In terms of Processing complexity and number of message exchange clustering algorithm should be efficient.
4) CHs are well-distributed over the network so that it cover entire network efficiently.
A. Clustering Objective and Advantages

In contrast to flat routing schemes, cluster-based routing scheme is having variety of advantages. Therefore, we summarize the advantages and objectives [9] of WSN clustering schemes are as follows:

**More Scalability:** Cluster-based routing scheme is more scalable as compare to flat topology. It is very easy to manage the events in the network.

**Data Aggregation/Fusion:** Data aggregation is the most popular method in which each cluster head perform the data aggregation, and sends it to the BS. It saves significant amount of energy.

**More Robustness:** Cluster-based routing schemes should be dealt with any network changes, unpredictable failure, node mobility etc. Routing schemes only have to do with these changes within individual cluster, in such a way robustness is achieved in entire network.

**Collision Avoidance:** Sensor network is divided into clusters and communications among nodes involves two modes i.e. intra-cluster and inter-cluster, for the purpose of data gathering and data transmissions respectively.

**Load Balancing:** Load balancing is an essential requirement in WSNs, aiming to increase the network lifetime. In cluster creation, approximately equal distribution of sensor nodes is usually considered. Further, CHs performs the intra-cluster management. Generally uniform clusters are adopted for increasing the network lifetime.

**Guarantee of Connectivity:** Due to the connectivity of each node, data is successfully delivered to the base station determines by the connectivity of each node along the path.

**Maximum Network Lifetime:** Network lifetime is most important requirement in WSNs, because limited power source of sensor nodes, transmission bandwidth and processing capability mainly for the applications of harsh environments.

B. Clustering properties

In cluster-based approaches, there are some characteristic for the cluster formation. These are related to the internal structure of the cluster. Some important characteristics [7] are as follows:

**Cluster count:** Cluster count is the number of clusters formed in a round. More number of cluster leads to small size cluster distribution, which is better in term of energy consumption.

**Cluster size:** Cluster size is the maximum path length among the member nodes from CH. Small size cluster is better in term of energy consumption because it minimizes transmission distance.

**Cluster Density:** Cluster density is defined as proportion of the number of cluster member in the cluster. In WSN, it is a big challenge to minimize the energy consumption in dense clusters.

**Message count:** Message count is the number of message transmission is requiring for CH selection. More number of message transmission leads to large amount of energy consumption for CH selection procedure.

**Stability:** If the cluster members are not fixed the clustering schemes are said to be adaptive. Otherwise, one can consider it as fixed because the cluster count is not varied throughout the clustering process. The fixed number of clusters increases the stability of a WSN.

**Intra-cluster topology:** It indicates the communication within the clusters either single hop or multi hop, from sensor node to sensor node or sensor node to CH. However, this communication also depends on the range of sensors.
Inter-cluster head connectivity: It indicates the capabilities of sensor nodes/CHs communication to BS.

4. Classification of Clustering Scheme

There are several ways to classify the clustering algorithms like centralized, distributed and hybrid routing algorithm [6]. In centralized clustering approach centralized architecture is used and there is fixed CH. In distributed approach, CH keep on changing based on some parameter such as residual energy, node degree etc. Reliability in WSN is improved by using Distributed approach. This paper elaborated some miscellaneous cluster-based routing schemes and their merits and limitation.

4.1. Low Energy Adaptive Clustering Hierarchy (LEACH)

LEACH was proposed by W. R. Heinzelman et al. [11]. It is first cluster-based distributed approach in WSN. The main aim of LEACH is to select sensor nodes as CHs by rotation, so that energy dissipation in communicating with the BS is spread to all sensor nodes in the network.

The LEACH operation is broken up into various rounds each round is divided into two phases, the set-up phase and the steady-state phase. Clusters are organized during the set-up phase and the data is delivered to the BS during steady-state phase. During the set-up phase, each node confirms whether or not to become CH for the current round. The nodes are selected based on selecting a random number between 0 & 1. The node becomes a CH for the current round if the number is lesser than the following threshold:

\[
T(n) = \frac{P}{1 - P(r \mod 1/P)} \quad \text{if } n \in G
\]  

In equation (1), P is the desired percentage of CHs, r is the current round and G is the set of nodes that have not been chosen as CHs in the last 1/P rounds. Cluster head rotation is performed to distribute the load among CHs. During the steady-state phase, the CHs perform the data aggregation and send this aggregated directly to the BS. LEACH uses a TDMA/CDMA MAC to reduce inter-cluster and intra-cluster collisions. After a certain time period the network goes back to the set-up phase again. In the literature, various alteration have been made to LEACH protocol, such as two level-LEACH (TL-LEACH) [12], Energy-LEACH (E-LEACH) [13], Weight-LEACH (W-LEACH) [14] etc.

The advantages of LEACH protocol [11] are as follows: (1) Each node have equal chance to be elected as a CHs, node that elected as a CH in current round cannot be elected as the CH again in further rounds. (2) Collisions can be avoided using TDMA schedule. (3) Excessive energy dissipation can be avoided by the cluster members in open and close communication interfaces with their time slots. However, there are some limitations with LEACH protocol [10,11] as follows: (1) LEACH uses single-hop communication. Therefore, it cannot be used in large networks. (2) CHs are elected only on the basis of probability, no other parameter are taking into consideration. (3) CHs are elected on the basis of probability. Therefore, there is a chance that the elected CHs are concentrated in one part of the network. (4) The dynamic clustering brings extra overheads.

4.2. Hybrid Energy Efficient Distributed (HEED)

HEED is proposed by O. Younis et al. [15]. It is energy efficient, multi-hop and distributed clustering protocol for WSNs. The main difference between HEED and LEACH is in election process of cluster head. CH election in HEED is not random, residual energy and communication-costs are the main parameters used during cluster formation. The main goals of HEED protocol [15] are:

1) Distribute energy consumption to increase the network lifetime.
2) Minimize energy during the CH election phase.
3) Minimize the control overhead of the network.
4) Even distribution of CHs throughout the networks.

Initially, 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:

\[
\text{CH}_{\text{prob}} = C_{\text{prob}} \left( \frac{E_{\text{residual}}}{E_{\text{max}}} \right)
\]

In equation (2), \( E_{\text{residual}} \) is the estimated current energy of the node and \( E_{\text{max}} \) is initial energy of a node or reference maximum energy, which is nearly same for all nodes present in the sensor network. The value of \( C_{\text{prob}} \) should not be allowed to drop down a certain threshold level that is selected to be inversely proportional to \( E_{\text{max}} \). The communication technique of HEED is similar to LEACH.

The advantages of HEED protocol [10] are as follows: (1) HEED is a distributed clustering routing technique. (2) HEED achieves load balancing and uniform CH distribution.

The limitations of HEED protocol [10, 9] are as follows: (1) In HEED, more number of CHs are generated due to that energy consumption is not balanced. (2) Similar to LEACH, large amount of overhead is created due to multiple rounds. (3) Cluster heads election process made additional overhead due to several iterations. (4) CHs, especially those are near the sink node, may dies earlier because of extra work load.

### 4.3. Threshold-Sensitive Energy Efficient Sensor Network (TEEN)

The TEEN was proposed by A. Manjeshwar et al. [16]. It is used to deal with sudden changes in the sensed attributes such as temperature. TEEN is a combination of hierarchical and data-centric approach. TEEN has a two-tier clustering topology shown in figure 1.

![Figure 1. The 2-tier Clustering Topology in TEEN](image)

It considers two thresholds into account, these are: hard threshold (HT) and soft threshold (ST). HT is used for the sensed attribute and ST is used for small changes in the value of the sensed attribute. In TEEN, every CH sends HT and ST values to its members. The HT and ST try to reduce amount of data communications.

The advantages of TEEN protocol [10] are as follows: (1) Data transmission can be controlled by varying two thresholds HT and ST. (2) It is suited to time-critical applications. However, there are some limitations with TEEN protocol [10, 9] as follows: (1) when thresholds are not satisfied, the node will no longer able to communicate, and if the node dies, the network will collapse and no longer be able to sense it. (2) If CHs are not in the range of one another, no communication is made and loss of data may occur. (3) It is not advisable for periodic reports Applications.

APTEEN introduced by A. Manjeshwar et al. [17]. It is an extension to TEEN and aims at both transmitting periodic data and reacting to time critical events. The structural design of APTEEN is similar to TEEN. In APTEEN firstly the clusters are formed by base station, now the CHs broadcast the attributes, the transmission schedule, count time (CT) and the threshold values to all nodes. All the nodes sense the environment continuously, if sensed value is equal to or beyond HT then transmission is permitted. If a node does not send data for a time interval equal to the CT, it must sense and transmit the data again. CHs achieve data aggregation in APTEEN in order to save energy.

The advantages of APTEEN protocol [9] are as follows: (1) APTEEN is suitable for both proactive and reactive applications. (2) By setting the CT interval and threshold values, APTEEN offers a lot of flexibility. For example energy consumption can be adjusted by changing the CT and the threshold values.

There are some limitations with APTEEN protocol [9] as follows: (1) Both TEEN and APTEEN having same drawbacks of additional overhead and complexity of cluster creation in multiple levels, setting the threshold values, and dealing with attribute-based naming of queries more difficult in APTEEN. (2) There is additional complexity required to implement the threshold values and the CT.

4.5. Power-Efficient Gathering In Sensor Information Systems (PEGASIS)

PEGASIS was proposed by S. Lindsey et al. [18]. It is an improved version of LEACH. In PEGASIS protocol, instead of organizing clusters it forms a chain from source to sink. Each node is executing a greedy algorithm in order to form a chain which is formed as farther nodes to nodes closer to the sink. Each node sends its sensed information to a close neighbor in the chain and data aggregation is performed at each node. Finally, all data are aggregated in one node i.e. leader node. Only leader node transmits the aggregated data to the sink. It increases the network lifetime. In, PEGASIS energy conservation is achieved in two ways:

1. The leader node (sink) receives at most two data messages.
2. The distance over which the data are transmitted to closest node is much smaller.

So, PEGASIS conserve energy by reducing the number of data messages at leader node (sink) [2]. The scheme of data transmission in PEGASIS is shown in Figure 2. In this figure, if node A5 is the leader, it will pass the token along the chain to node A1 at first. Then, node A5 will forward its data toward node A6. After node A5 receives data from node A4, it will pass the token to node A4, and node A1 will forward its data towards node A8 with data aggregation taking place along the chain.

The advantages of PEGASIS [10,19] are as follows: (1) PEGASIS reduces the overhead due to dynamic cluster formation. (2) It reduces the number of data transmissions due to the chain of data aggregation. (3) The energy load is distributed uniformly in the network because all nodes take turns acting as leader.

![Figure 2. The Token Passing Scheme in PEGASIS](image)

There are some limitations with PEGASIS [10,19] as follows: (1) PEGASIS is not suitable for networks for dynamic topologies. (2) Communication has very long delays,
which leads to bottle neck into nodes. (3) All the nodes have to know globally about the network due to that network is not very scalable.

4.6. Energy Efficient Clustering Scheme (EECS)

EECS was proposed by M. Ye et al. [20]. It is suitable for periodic data gathering applications. It is an extension of LEACH and differs from LEACH in the cluster setup stage. The main idea of EECS is to elect CHs with more residual energy. EECS divides the network into clusters and uses direct single-hop communication between the CH and BS. EECS creates uniform CHs across the network.

The advantages of EECS protocol [10] are as follows: (1) EECS constructs a more balanced network in terms of energy consumption and communication load. (2) As some CH will have to communicate over long distances with the BS, they consume more energy to communicate, to solve this problem, it uses dynamic sizing of clusters.

The limitations of EECS [10] are as follows: (1) A lot of energy is used to communicate between CH and BS due to single-hop communications. (2) It uses global information for communication, which made an additional overhead. (3) Overhead complexity is created during CH election process due to competition among nodes.

4.7. Concentric Clustering Scheme (CCS)

The CCS was proposed by S. Jung et al. [21]. It is an extension of PEGASIS and reduces the energy consumption. The main concept of CCS is to take into consideration the BS location to increase the network lifetime. In CCS, the network is divided into concentric circular tracks, those are treated as different clusters, and level is given to each circular track. The track closed to the BS is assigned as level 1, and the level number increases as the distance from the BS increases. Similar to PEGASIS, chains are constructed inside the track. In each level, one node is elected as a CH and data transmission is based on the PEGASIS scheme. All the nodes in each level transmit their data to the nearest node along the chain. The receiving node fuses its own data to the data received and transmits this data to the next node. The CH node in level-n send out data to the CH node in level (n-1) and this process persist until sending data to the BS is ended. Data aggregation can be done at every CH. The data transmission scheme in CCS is shown in Figure 3.

Some limitations of CCS over PEGASIS protocol [10] are as follows: (1) In CCS, energy consumption is reduced due to data transmission distance from CH to BS is reduced. (2) CCS also saves energy by dividing the network into concentric clusters.

![Figure 3. The Data Transmission Scheme in CCS](image-url)
next hop is based on the location of nodes, which may cause the energy of a CH to dissipate rapidly on the path among CHs.

**4.8. Track-Sector Clustering (TSC)**

TSC was proposed by N. Gautam et al. [22]. The network is partitioned into concentric circular tracks and triangular sectors. The division of tracks and sectors saves energy because it minimizes redundant data transmission and provides the shortest distance between CH and BS. A cluster in TSC is an area under a curved strip formed by the intersection of a circular track and a triangular sector.

The TSC process execution is divided into four phases, these are: track setup, sector setup and CH election, chain construction and data transmission. In track setup phase, the BS sets the concentric circular tracks and keeping itself at the geometric center of the circular tracks. Each node assigned a particular track and each track is designated a level. The total number of tracks depends on node density and the location of the BS. During the sector setup and CH election, sectors are formed and initially few CHs are elected by the BS such that every sector has a CH. For the chain construction, chains are constructed by the intersection of tracks and sectors within each cluster. During the data transmission, member nodes transmit and receive data to the neighbor node within a cluster. Finally the CH in each cluster performs fusion and transmits it to the CH in the lower level track and finally data sent to the BS through multi hops.

The advantage of TCS over CCS and PEGASIS [10] is: It has tracks and sectors, TSC reduces redundant data transmission in the network. Some limitations of TSC [10] are as follows: (1) Distribution of node in each level is unbalanced. Therefore, the lower level nodes deplete their energy first. (2) Residual energy is not considered for CHs selection, so this scheme may lead to unbalanced energy consumption across the network.

*Clustering issues and challenges*

There are several issues [10, 19] to be tackled for the efficient use of clustering routing protocol. Some open issues are: calculation and selection of CHs, network topology, Fault tolerance and redundancy management.

**5. Comparison of Cluster-based Routing Schemes**

Table 1. summarizes the comparison between cluster-based routing schemes on the basis of scalability, cluster stability, energy efficiency, load balancing, algorithm complexity and delivery delay. LEACH, TEEN, APTEEN and PEGASIS have similar features and fixed architecture in some extent.

This paper presents Routing schemes which are of miscellaneous category in terms of their architecture and methodology used in cluster formation and cluster head selection. As if we consider the energy and network life time is the main parameters then the performance of APTEEN lies between TEEN and LEACH with respect to energy consumption and prolonged network. TEEN only transmits time-critical data, while APTEEN transmits periodic data. In this way APTEEN is also better than LEACH. PEGASIS avoids the excessive cluster overhead but introduced the excessive delay due to chain based architecture.
this paper shows the comparison of different clustering protocols on the basis of various performance measures such as scalability, energy efficiency and load balancing etc. It is observed that the different clustering-based routing schemes mentioned above can be used to improve the performance in terms of energy efficiency, network lifetime, load balancing and other metrics. This paper is based on a comparative study of existing clustering-based routing protocols in WSN, which provides the scope for the researcher on optimized and energy-saving sensor networks.

References


Table 1. Comparison of different Cluster-based Routing Schemes

<table>
<thead>
<tr>
<th>Schemes Name</th>
<th>Scalability</th>
<th>Cluster Stability</th>
<th>Energy Efficiency</th>
<th>Load Balancing</th>
<th>Algorithm Complexity</th>
<th>Delivery Delay</th>
</tr>
</thead>
<tbody>
<tr>
<td>LEACH</td>
<td>Very low</td>
<td>Moderate</td>
<td>Very low</td>
<td>Moderate</td>
<td>Low</td>
<td>Very small</td>
</tr>
<tr>
<td>HEED</td>
<td>Moderate</td>
<td>High</td>
<td>Moderate</td>
<td>Moderate</td>
<td>Moderate</td>
<td>Moderate</td>
</tr>
<tr>
<td>EECS</td>
<td>Low</td>
<td>Moderate</td>
<td>Moderate</td>
<td>Moderate</td>
<td>Very high</td>
<td>Moderate</td>
</tr>
<tr>
<td>TEEN</td>
<td>Low</td>
<td>Moderate</td>
<td>Very high</td>
<td>Good</td>
<td>High</td>
<td>Small</td>
</tr>
<tr>
<td>APTEEN</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Moderate</td>
<td>Very high</td>
<td>Small</td>
</tr>
<tr>
<td>PEGASIS</td>
<td>Very low</td>
<td>Low</td>
<td>High</td>
<td>High</td>
<td>Very large</td>
<td>Moderate</td>
</tr>
<tr>
<td>CCS</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Very bad</td>
<td>Moderate</td>
<td>Large</td>
</tr>
<tr>
<td>TSC</td>
<td>Moderate</td>
<td>Moderate</td>
<td>Moderate</td>
<td>Bad</td>
<td>Moderate</td>
<td>Moderate</td>
</tr>
</tbody>
</table>
Authors

Pooja Mishra, she has received her B.E. in Information Technology from Madhav Institute of Technology and Science, Gwalior, India in 2013. Currently she is pursuing M. Tech. in Information Technology from Madhav Institute of Technology and Science, Gwalior. Her area of interest includes wireless sensor networks, wireless communication.

Sanjiv Sharma, he currently works as an Asst. professor in Department of Computer science Engineering and Information Technology at Madhav Institute of Technology and Science, Gwalior. He has 9 years of teaching and research experience. He has received his B.E. from Pt. Ravishankar University, Raipur, and M. Tech. from RGPV, Bhopal and Ph. D from Bansthali Vidhya Peeth, Jaipur. His area of research is Network security, Data Mining and Ad-hoc and wireless sensor Networks. and Science, Gwalior. Her area of interest includes wireless sensor networks, wireless communication.