A Survey on Prominent Clustering Schemes in Wireless Sensor Networks

Abey Abraham  
Asst. Professor,  
Department of Information Technology,  
Rajagiri School of Engineering and Technology,  
Kochi, India

Tina Sebastian  
PG student,  
Department of Information Technology,  
Rajagiri School of Engineering and Technology,  
Kochi, India

ABSTRACT
There is an increased interest in the use of wireless sensor networks (WSNs) for the past few years. Energy constraint is a critical problem to be considered. Clustering is introduced in WSNs because of its network scalability, energy-saving attributes and network topology stabilities. Generally clustering can be classified into three methodologies: Centralized clustering, Distributed clustering, Hybrid clustering. Clustering is becoming an active branch of routing technology in WSNs. This paper presents a comprehensive and fine grained survey on various clustering schemes in WSN. A few prominent WSN clustering routing protocols are analyzed and compared these different approaches based on our taxonomy and several significant metrics.

General Terms  
Sensor Networks, Routing Protocols

Keywords  
Clustering, ClusterHead Energy Efficiency, Wireless Sensor Network

1. INTRODUCTION
A Wireless sensor network is formed by spatially distributing low powered small sensor nodes communicating among themselves using radio signals and deployed randomly or manually in an unattended environment having limitations in power, sensing and processing capabilities. Sensor nodes are available in large numbers at a low cost to be employed in a wide range of applications[1]. The sensor nodes are also capable of performing other functions such as data processing and routing. Grouping sensor nodes into clusters has been widely adopted by the research community to satisfy the above objectives and generally achieve high energy efficiency and prolong network lifetime in large-scale WSN environments. Generally clustering can be classified into three methodologies[7]. First method is centralized clustering where the base station configures the entire network into clusters, second method is distributed clustering where the sensor nodes configure themselves into clusters and third method is Hybrid clustering which is formed as the resulting configuration of the above two methods[7].

2. CLUSTERING SCHEMES
Clustering can be classified into three methodologies[7]. First method is centralized clustering where the base station will configure the entire network into clusters, second method is distributed clustering where the sensor nodes configure themselves into clusters and third method is Hybrid clustering which is formed as the resulting configuration of the above two methods[7].

2.1 Centralized Clustering
In centralized methods, a sink or CH requires global information of the network or the cluster to control the network or the cluster. The efficiency is limited in large-scale networks where collecting all the necessary information at the central authority is both time and energy consuming[4].

2.1.1 Firefly Algorithm
Firefly algorithm was introduced by Dr. Xin She yang at Cambridge University in 2007, modeled after the flashing behavior of fireflies[8]. The aim of Firefly Algorithm is to find the particle position that results in the best evaluation of a given fitness function. There are three main rules:

- Fireflies are unisex. That is a firefly will be attracted by other fireflies regardless of their sex.
- The attractiveness of a Firefly is directly proportional to its brightness and decreases as the distance increases.
- The objective function results the brightness of a firefly.

Advantages and disadvantages of Firefly algorithm: Firefly algorithm is a favorable optimization tool due to the effect of the attractiveness function. It not only includes the self improving process with the current space, but it also includes the improvement among its own space from the previous stages[7]. Firefly algorithm has some disadvantages like getting trapped into several local optima. It sometimes performs local search as well and sometimes is unable to completely get rid of them. Parameters are fixed and they do not change with time. Firefly does not memorize or remember any history of better situation, and they may end up missing their situations[7].
Algorithm 1: Firefly Algorithm

2.1.2 Jumper Firefly Algorithm

Mahdi Bidar and Hamidreza Rashidy Kanan developed a new algorithm [9] based on firefly algorithm to improve the performance of the agents in determining more appropriate solutions by modifying them, thereby the probability of finding the optimal solution can be increased. A Status Table is used which records and observes all the details of the Fireflies behavior. Status table helps to indicate the agents that have to be changed in their situations by jumping into new situations. Table 1 shows the Status Table used in this algorithm[9].

Algorithm 2: Jumper Firefly Algorithm

Advantages of Jumper Firefly algorithm: Jumper Firefly Algorithm prolongs network life time. It gives better network partitioning with minimum intra-cluster distance[7]. Energy consumed by all the nodes for communication can be reduced.

Table 1. Status Table

<table>
<thead>
<tr>
<th>Parameter</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>...</th>
<th>n</th>
</tr>
</thead>
</table>

Cluster setup using Jumper Firefly algorithm is similar to that using Firefly algorithm but Jumper Firefly algorithm is implemented at the base station instead of Firefly algorithm.

2.1.3 BCDCP

Base-Station Controlled Dynamic Clustering Protocol (BCDCP) was introduced by Muruganathan et al. [8]. BS receives information about residual energy from every node. Based on this, it computes the average energy level of all the nodes in the network, and then chooses a set of nodes whose energy levels are above the average value[6]. Only the nodes from the chosen set can be elected CHs for the current round. Based on the chosen set, the BS computes the number of clusters and performs the task of clustering.

Advantages and disadvantages of BCDCP: BCDCP solves the problem of CH distribution and ensures similar power dissipation of CHs. There are some drawbacks: Limited scalability and robust to large networks. Increased design complexity and energy consumption of nodes. Due to the single-hop routing scheme, it is not appropriate for long-distance communications[12]. BCDCP is not adaptive to applications in large-scale networks. It is not suitable for reactive networks.

2.2 Distributed Clustering

In Distributed approaches, a sensor node becomes a CH or joins a formed cluster on its own initiative without global information of the network or the cluster.

Distributed algorithms are more suitable for large-scale networks. In such approaches, a node decides to join a cluster or become a CH based on information obtained only from its one-hop neighbors[3].

2.2.1 LEACH

Low-Energy Adaptive Clustering Hierarchy (LEACH), was proposed by Heinzelman et al. [16]. The main objective of LEACH is to select sensor nodes as CHs by rotation, so the high energy dissipation in communicating with the BS is spread to all sensor nodes in the network. There are several rounds and each round is separated into two phases, set-up phase and steady-state phase.

Advantages and disadvantages of LEACH: Each node can equally share the load imposed upon CHs. Disadvantages include: Long-range communications directly from CHs to the BS will lead to too much energy consumption. LEACH cannot ensure real load balancing in the case of sensor nodes with different amounts of initial energy. Since CH election is performed in terms of probabilities, it is difficult for the predetermined CHs to be uniformly distributed throughout the network[12].
2.2.2 **HEED**

Hybrid Energy-Efficient Distributed clustering (HEED) [8], was introduced by Younis and Fahmy. HEED does not select nodes as CHs randomly. The manner of cluster construction is performed based on the combination of two parameters, node’s residual energy and intra-cluster communication cost[12].

**Advantages and disadvantages of HEED:** Low power levels of clusters promote an increase in spatial reuse while high power levels of clusters are required for inter-cluster communication. There is uniform CH distribution across the network and load balancing. Multi-hop communication between CHs and the BS promote more energy conservation and scalability[5]. Disadvantages are: The use of tentative CHs that do not become final CHs leave some uncovered nodes. They are forced to become a CH and these forced CHs may be in range of other CHs or may not have any member associated with them. More CHs will be generated[10]. Overhead causes noticeable energy dissipation which results in decreasing the network lifetime[7].

2.2.3 **DWEHC**

Distributed Weight-based Energy-efficient Hierarchical Clustering protocol (DWEHC) was proposed by Ding et al. [11]. The main objective of DWEHC is to improve HEED by building balanced cluster sizes and optimize intra-cluster topology using location awareness of the nodes. Every node implements DWEHC individually and the algorithm ends after several iterations. Locally calculated parameter weight is defined for CH election in DWEHC. Intra-cluster communication is performed by TDMA.

2.3 **Hybrid Clustering**

Hybrid schemes are composed of centralized and distributed approaches. Distributed approaches are used for coordination between CHs and centralized for CHs to build individual clusters.

### 3. COMPARISON

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Firefly Algorithm</th>
<th>Jumper Firefly Algorithm</th>
<th>BCDCP</th>
<th>LEACH</th>
<th>HEED</th>
<th>DWEHC</th>
<th>RCC</th>
<th>GROUP</th>
<th>S-WEB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clustering</td>
<td>Centralized</td>
<td>Centralized</td>
<td>Centralized</td>
<td>Distributed</td>
<td>Distributed</td>
<td>Distributed</td>
<td>Hybrid</td>
<td>Hybrid</td>
<td>Hybrid</td>
</tr>
<tr>
<td>Cluster Stability</td>
<td>Less Stable</td>
<td>Stable</td>
<td>High</td>
<td>Moderate</td>
<td>High</td>
<td>High</td>
<td>Provisioned</td>
<td>Stable</td>
<td>Stable</td>
</tr>
<tr>
<td>Cluster Size</td>
<td>Even</td>
<td>Even</td>
<td>Even</td>
<td>Even</td>
<td>Even</td>
<td>Even</td>
<td>Even</td>
<td>Even</td>
<td>Even</td>
</tr>
<tr>
<td>Cluster Count</td>
<td>Variable</td>
<td>Variable</td>
<td>Variable</td>
<td>Variable</td>
<td>Variable</td>
<td>Variable</td>
<td>Variable</td>
<td>Controlled</td>
<td>Variable</td>
</tr>
<tr>
<td>Intra Cluster Topology</td>
<td>Minimized</td>
<td>Minimized</td>
<td>1 Hop</td>
<td>1 Hop</td>
<td>1 Hop</td>
<td>k Hop</td>
<td>k Hop</td>
<td>k Hop</td>
<td>Minimized</td>
</tr>
<tr>
<td>Inter Cluster Topology</td>
<td>Single hop Multiple hop</td>
<td>Single hop Multiple hop</td>
<td>Single hop Multiple hop</td>
<td>Direct Link</td>
<td>Multihop Hierarchical</td>
<td>Single hop</td>
<td>Direct Link</td>
<td>Multihop Hierarchical</td>
<td>Single hop Multi hop</td>
</tr>
<tr>
<td>Cluster Overlap</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Control Message</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 3. Cluster Head Characteristics
4. CONCLUSION

Wireless sensor networks have attracted significant attention over the past few years, and can be employed in a wide spectrum of applications. The design of effective, robust, and scalable routing protocols for WSNs is a challenging task. Significant efforts have been made in addressing the techniques to design effective and efficient clustering routing protocols for WSNs in the past few years. In this paper, we have presented survey on different clustering schemes employed in WSN namely Centralized, Distributed and Hybrid. We have seen few clustering protocols based on them. We have systematically analyzed them in detail and compared these different approaches based on our taxonomy and some primary metrics.
5. REFERENCES


