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[2]. In centralized, a sink or CH requires global information of the network or the cluster to control the network or the cluster. In distributed, a sensor node is able to become a CH or to join a formed cluster on its own initiative without global information of the network or the cluster. Hybrid schemes are composed of centralized and distributed approaches[5]. Here, distributed approaches are used for coordination between CHs, and centralized approaches are performed for CHs to build individual clusters[2].The remainder of this paper is organized as follows: Section 2 analyzes few prominent WSN clustering routing protocols. Section 3 compares different approaches. Finally Section 4 summarizes and concludes the paper.

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].

```
Data:
  • Generate S particles to contain K randomly selected cluster heads.
  • Map the randomly generated positions with the closest (x,y) coordinates.
Result: Cluster heads positions are obtained

while(t< Max Generation)
Map the positions with the closest (x,y) coordinates
Evaluate the cost function
for i=1 to n (all n fireflies)
for j=1 to n (all n fireflies)
if (I_j > I_i
Update the particles positions;
Limit the change in the particle's position value;
end
end
end
Rank the fire flies and find the current best;
end
Post process the results;
End procedure;
```

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].

```
Data:
  • Generate S particles to contain K randomly selected cluster heads.
  • Define the user defined constants like η, Ω, β...
  • Map the randomly generated positions with the closest (x,y) coordinates.
  • Create the Status table
Result: Cluster heads positions are obtained

while(t< Max Generation)
if(any firefly is in hazard)
Put the firefly in new position stochastically;
Update the Status table;
end
Map the positions with the closest (x,y) coordinates
Evaluate the cost function
for i=1 to n (all n fireflies)
for j=1 to n (all n fireflies)
if(I_j > I_i)
Update the particles positions;
Limit the change in the particle's position value;
Update the Status table;
end
end
end
Rank the fire flies and find the current best;
end
Post process the results;
End procedure;
```

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

Firefly	1	2	3	…	n
Parameter					
Situation					
Fitness					
Worst					
Qualification					

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-range 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.

2.3.1 RCC

Random Competition Based Clustering [13] was designed for mobile ad-hoc networks. RCC focuses on cluster stability to support mobile nodes. It applies the First Declaration Wins rule, where any node can govern the rest of the nodes in its radio coverage if it is the first to claim a CH. After hearing the claim which is broadcasted by the first node, neighboring nodes join its cluster as member nodes and they give up their right to be a CH. Every CH in the network broadcast a CH claim packet to maintain clusters. Time delay between broadcasting a claim packet and receiving it causes a conflict[15]. To avoid this RCC employs a random timer and uses the node ID for arbitration. During this random time if it receives a broadcast message carrying CH claim packet from another node, it ceases transmission of its CH claim. If the conflict exists, node having lower ID will become CH.

2.3.2 GROUP

Grid-clustering ROUting Protocol provides efficient and scalable packet routing for large-scale WSNs. The primary sink dynamically and randomly builds the cluster grid, where CHs are arranged in a grid manner[12]. Forwarding of data queries from the sink to source node is propagated from the Grid Seed (GS) to its CHs. For location unaware data query the query is passed from the central most sink in the network to its nearest CH[16]. It will then broadcast the message to neighbouring CHs. If it is location aware, then the request will be sent down the chain of CHs towards the particular region using unicast packets.

2.3.3 S-WEB

Sensor Web or S-WEB divides the sensing field into clusters bordered by two arcs of two adjacent concentric circles and two adjacent radii from the BS. Clusters are identified by angle order (β) and the order of Signal Strength threshold (δ)[14]. BS in S-WEB will send beacon signals for every α degree angle, one at a time. Sensors that receive the beacons at time slot i will measure their signal strength to determine their relative distances to the BS.

3. COMPARISON

Table 2. Cluster Characteristics

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

Table 3. Cluster Head Characteristics

Paramete r	Firefly Algorithm	Jumper Firefly Algorithm	BCDCP	LEACH	HEED	DWEHC	RCC	GROUP	S-WEB
CH Selection	Random	Random	Random	Random	Random	Weight Based	Random	Connectivi ty	Residual Energy
CH Rotation	Yes	Yes	Yes	Yes	Yes	No	No	No	Yes
Mobility	Limited	Mobile	Stationary	Limited	Limited	No	Yes	No	Yes
Existence	Cluster Head Based	Cluster Head Based	Cluster Head Based	Cluster Head Based	Cluster Head Based	Cluster Head Based	Cluster Head Based	Cluster Head Based	Cluster Head Based
Role	Relay Aggregatio n	Relay Aggregation	Relay Aggregatio n	Relay Aggregatio n	Relay Aggregatio n	Relay Aggregatio n	Relay	Relay	Relay Aggregation
Node Type	Sensor	Sensor	Sensor	Sensor	Sensor	Sensor	Sensor	Sensor	Sensor
Paramete r for CH election	Adaptive	Adaptive	Adaptive	Adaptive	Adaptive	Adaptive	Adaptiv e	Adaptive	Adaptive

Table 4. Clustering Characteristics

Parameter	Firefly Algorit hm	Jumper Firefly Algorithm	BCDCP	LEACH	HEED	DWEHC	RCC	GROUP	S-WEB
Energy Efficiency	Efficien t	Optimized	Very low	Very low	Moderate	Very high	N/A	Efficient	Efficie nt
Scalability	Moder ate	Good	Very low	Very low	Moderate	Moderate	Good	Good	Good
Load Balancing	Good	Good	Good	Moderate	Moderate	Very good	Good	Good	Good
Algorithm Stages	Cluster Constr uction	Cluster Constructi on	Cluster Constructio n	Cluster Construction	Cluster Construction	Cluster Constructio n	Cluster Constr uction	Cluster Construc tion	Cluste r Constr uction
Algorithm Complexity	Moder ate	Moderate	Very high	Low	Moderate	Moderate	Variabl e	Moderat e	Mode rate
Time Complexity	Variabl e	Variable	High	Constant	Constant	Constant	Variabl e	Variable	Variab le
Execution nature	Iterativ e	Iterative	Iterative	Probabilistic	Iterative	Iterative	Iterativ e	Iterative	Proba bilistic
Convergence Time	Variabl e	Variable	Constant	Constant	Constant	Constant	Variabl e	Variable	Variab le
Delivery Delay	Small	Small	Small	Very small	Moderate	Moderate	Mediu m	Medium	Mediu m
Clock Synchronizati on	No	No	No	No	No	No	No	No	No

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

[1] S. W. S. Y. C. E. Akyildiz, I. F., "Wireless sensor networks: a Survey", in Elsevier Sci B.V. Computer networks 38, vol 1, no.4, 2002, pp. 393–422.

[2] W. R. Heinzelman, A. Chandrakasan, and B. Hari, "Energy Efficient Communication Protocol for Wireless Microsensor Networks," in System Sciences, 2000. Proceedings of the 33rd Annual Hawaii International Conference on. IEEE, 2000.

[3] R. R. Tillett, Jason and F. Sahin, "Cluster-head identification in ad hoc sensor networks using particle swarm optimization," in Personal Wireless Communications, 2002 IEEE International Conference on. IEEE, 2002, (pp. 201–205).

[4] W. R. Heinzelman, A. Chandrakasan, and B. Hari, "an Application Specific Protocol Architecture for Wireless Microsensor Networks," in Wireless Communications, IEEE Transactions on 1 no 4 (2002). IEEE, 2002, pp. 660–670.

[5] N. A. Latiff, C. Tsimenidis, and B. Sharif, "Energy Aware Clustering for Wireless Sensor Networks using Particle Swarm Optimization," in Personal, Indoor and Mobile Radio Communications. PIMRC 2007. IEEE 18th International Symposium on.IEEE, 2007, pp. 1–5.

[6] V. Kumar , S. Jain, S. Tiwarietal, "Energy Efficient Clustering Algorithms in Wireless Sensor Networks: A survey," IJCSI International Journal of Computer Science Issues, vol. 8, no. 5, pp. 1694–0814, 2011.

[7] Prof. N.V.S.N Sarma, Mahesh Gopi, "Energy Efficient Clustering using Jumper Firefly Algorithm in Wireless Sensor Networks," International Journal of Engineering Trends and Technology (IJETT) , vol 10, pp. 525-532, April 2014.

[8] A. Hashmi, N. Goel, S. Goel, and D. Gupta, "Firefly Algorithm for Unconstrained Optimization," IOSR Journal of Computer Engineering (IOSR-JCE), vol. 11, pp. 75–78, May-June 2013.

[9] M. Bidar and H. R. Kanan, "Jumper firefly algorithm," International Conference on Computer and Knowledge Engineering (ICCKE-2013), pp. 267–271, Oct-Nov 2013.

[10] J. Wan, D. Yuan, X. Xu, "A review of cluster formation mechanism for clustering routing protocols," in: Proceedings of 2008 11th IEEE International Conference on Communication Technology (ICCT 2008), Hangzhou, China, pp.611-616, November 2008.

[11] J. Al-Karaki, A. Kamal, "Routing techniques in wireless sensor networks: a survey," IEEE Wireless Communications, vol. 11,pp. 6-28, June 2004.

[12] Xuxun Liu, "A survey on Clustering Routing protocols in wireless sensor networks," www.mdpi.com/journal/sensors 2012, 12, 11113-11153.

[13] J. Yu, P. Chong, "A survey of clustering schemes for mobile ad hoc networks," IEEE Communications Surveys & Tutorials, vol. 7,pp. 32- 48, January 2005.

[14] A. Abbasi, M. Younis, "A survey on clustering algorithms for wireless sensor networks," Computer Communications, vol. 30,pp. 2826-2841, October 2007.

[15] Y. Li, M. Thai and W. Wu, "Topology control for wireless sensor networks," Wireless Sensor Networks and Applications, Heidelberg: Springer, 2008, pp.113-147.

[16] W. Cheng and H. Shi, "AEEC: an Adaptive Energy Efficient Clustering Algorithm in Sensor Networks", in Industrial Electronics and Applications. ICIEA. 4th IEEE Conference on. IEEE, 2009, pp. 3950–3954.