



A survey of Clustering Technique for Saving the Energy in Wireless Sensor Networks

S.Janakiraman¹, V.Sivakumar², K.Maharajan³

¹Department of Banking Tech., Pondicherry University, Puducherry, India

^{2,3}Department of CSE, National Engg. College, Kovilpatti, Chennai, Tamil Nadu, India

Abstract: *Wireless Sensor Networks (WSNs) consist of small nodes capabilities of sensing, communications and computation. Many routing, data dissemination and power management protocols have been specially designed for WSNs where energy awareness is an essential design issue. In sensor networks routing protocols which might differ depending upon the application and network architecture. In this paper, we present a survey of the state-of-the-art Hierarchical routing in WSNs. We first outline the design challenges for routing protocols in WSNs followed by a comprehensive survey of different clustering technique in Hierarchical routing. We also highlight the performance issues and advantages of each technique. The paper concludes with possible future research areas.*

Keywords: *energy efficiency, hierarchical routing, clustering*

I. INTRODUCTION

A Sensor networks are usually inexpensive and therefore WSNs may be composed of a huge number of sensor nodes, which themselves are deployed inside and/or around the phenomenon. WSNs are based on low-cost processors, resulting in limited energy budget and restricted memory space. In many applications, it is expected that the sensor node last for a long time because in most of the cases these networks are used in remote areas and recharging and/or replacing power supply units is considered difficult or prohibitive due to hazardous and inaccessible places where they are supposed to operate. The paper is organized as follows: Routing challenges and designing of WSN is described in the Section II. Section III reviews previous survey articles on WSN routing with clustering concept. And finally Section IV concludes the paper with proposed future directions.

II. ROUTING CHALLENGES AND DESIGN ISSUES

Wireless sensor networks have several restrictions, e.g., limited energy, computing power, and bandwidth of the wireless links connecting sensor nodes. One of the main design goals of WSNs is to carry out data communication while trying to prolong the lifetime of the network and prevent connectivity degradation by employing aggressive energy management techniques. The design of routing protocols in WSNs is influenced by many challenging factors. These factors must be overcome before efficient communication can be achieved in WSNs. A careful definition of protocols and algorithms for efficient communication has become one of the most important issues in WSNs in order to improve their lifetime. Depending on the network structure, routing in WSNs can be divided into:

- a) Flat-based routing
- b) Hierarchical-based routing
- c) Location-based routing

There are varieties of different routing protocols has been designed, On the basis of topologies routing protocols may lay on above network types. Among all topologies routing, hierarchal routing protocol technique is more popular regarding the power saving of sensor nodes. In this technique sensor nodes are formed by several clusters. Cluster is responsible to transfer the data to the sink. Clusters communication works on the basis of cluster leader which can be known as cluster head. Communication with sink can be done with the help of cluster head, they collect data from neighboring nodes and send it to another cluster head, which is responsible for any other cluster, this mechanism continuous until the data reaches to the sink. Basically, the Protocol must be energy efficient and must try to reduce the following issues related to energy consumption phenomena:

- **Packet collision:** When one node receives more than one packet at the same instant, it is considered that a packet collision occurred. Therefore, all packets must be discarded and transmitted again.
- **Overhearing:** When some node receives packets that are addressed to another sensor node overhearing occurred.
- **Control packet overhead:** The use of control packets in order to coordinate the WSN must be minimized.

- **Idle listening:** Idle listening occurs when some node is in the listening mode of a channel that is not being use
- **Over-emitting:** Over-emitting is caused when the message delivery fails due to the destination node's inactivity.
- **Latency:** Refers to the shortest total delay at the intermediate nodes or channels in transmitting a unique packet from all sending sources to a destination, which includes propagation, queuing and routing delay.

III. ENERGY EFFICIENT HIERARCHICAL CLUSTERING PROTOCOLS

A sensor is a battery-driven device and it not possible changing the battery frequently in the sensing field. However, every operation like sensing, transmitting and receiving operation a sensor consumes a certain amount of energy. To maximize the network lifetime, a sensor network protocol must be energy-efficient. The best way to energy-efficiently is to balance energy consumption among all the sensors in the network. For that, sensors which has more energetic it must have more chance to be active, and others are in passive mode. The past energy saving protocols as follows

1. LEACH (Low-energy adaptive clustering hierarchy) proposed in [1] is a hierarchical protocol in which most nodes transmit to cluster heads, and the cluster heads aggregate and compress the data and forward it to the base station(sink). Each node uses a stochastic algorithm at each round to determine whether it will become a cluster head in this round. LEACH assumes that each node has a radio powerful enough to directly reach the base station or the nearest cluster head, but that using this radio at full power all the time would waste energy. Nodes that have been cluster heads cannot become cluster heads again for P rounds, where P is the desired percentage of cluster heads. Thereafter, each node has a $1/P$ probability of becoming a cluster head in each round. At the end of each round, each node that is not a cluster head selects the closest cluster head and joins that cluster. The cluster head then creates a schedule for each node in its cluster to transmit its data. All nodes that are not cluster heads only communicate with the cluster head in a TDMA fashion, according to the schedule created by the cluster head. They do so using the minimum energy needed to reach the cluster head, and only need to keep their radios on during their time slot. This protocol has increased the lifetime of the network by nearly 50% of the original lifetime of the network.

2) M-LEACH (Multi-hop-LEACH) proposed in [2] improves on LEACH by Each CH directly communicates with BS no matter the distance between CH and BS. It will consume lot of its energy if the distance is far. On the other hand, Multihop-LEACH protocol selects optimal path between the CH and the BS through other CHs and use these CHs as a relay station to transmit data over through them. First, multi-hop communication is adopted among CHs. Then, according to the selected optimal path, these CHs transmit data to the corresponding CH which is nearest to BS. Finally, this CH sends data to BS. M-LEACH protocol is almost the same as LEACH protocol, only makes communication mode from single hop to multi-hop between CHs and BS.

3) LEACH-C (Centralized LEACH) proposed in [3] runs a centralized cluster formation algorithm so that the cluster heads can be distributed uniformly throughout the network. Though it reduces energy consumption over its LEACH siblings by balancing the distribution of cluster heads throughout the whole network, LEACH-C requires location information of the nodes usually provided by GPS devices which makes the protocol less robust and introduces some extra cost per node.

4) V-LEACH (Version LEACH) protocol proposed in [4] the cluster contains; CH (responsible only for sending data that is received from the cluster members to the BS), vice-CH (the node that will become a CH of the cluster in case of CH dies), cluster nodes (gathering data from environment and send it to the CH). In the original leach, the CH is always on receiving data from cluster members, aggregate these data and then send it to the BS that might be located far away from it. The CH will die earlier than the other nodes in the cluster because of its operation of receiving, sending and overhearing. When the CH die, the cluster will become useless because the data gathered by cluster nodes will never reach the base station.

5) TEEN (Threshold Sensitive Energy Efficient Network Protocol) [5] combines both data-centric and hierarchical approach. Instead of having a flat clustering of nodes like LEACH or PEGASIS, TEEN creates multi-level hierarchy of nodes by grouping closer nodes into clusters on the first level, then on the second level and so on until the sink node is reached. TEEN tries to reduce energy consumption by reducing the number of transmissions with the help of two thresholds as broadcast by the cluster heads after the clusters are formed – hard and soft thresholds. Hard threshold refers to the minimum possible attribute value that causes a sensor node to switch on its transmitter and transmit the sensed data to the cluster head, whereas a node sensing an attribute value at or beyond the hard threshold will transmit the sensed data only when the attribute value changes by an amount equal to or greater than the soft threshold. Although TEEN is quite energy-efficient, it is not useful for applications that require periodic reporting of the sensed data. Most importantly, the forming of multi-level hierarchy, implementing different thresholds and attribute-based naming impose huge overhead and complexity to this approach.

6) APTEEN (Adaptive TEEN) An extension to TEEN was proposed in [6] called Adaptive Threshold Sensitive Energy Efficient Network Protocol which overcomes the limitation of TEEN by incorporating the periodic reporting of sensed data along with the reactive response. The architecture of APTEEN which remains the same as TEEN as a result of which APTEEN inherits the drawbacks of TEEN as well.

7) DECSA (Distance-Energy Cluster Structure Algorithm) proposed protocol [7] is based on the classic clustering routing algorithm LEACH, it considering both the distance and residual energy of nodes. It improves the process of cluster head selecting and the process of cluster forming. It reduces the adverse effect on the energy consumption of the cluster head, resulting from the non-uniform distribution of nodes in network and avoided the direct communication between the Base Sensor and cluster head, which has low energy and far away from Base Sensor. An algorithm effectively balances the energy consumption, prolongs the lifetime, reduces the energy consumption and has a better performance than the original LEACH protocol.

8) EDDEEC (Enhanced Developed Distributed Energy-Efficient Clustering for Heterogeneous Wireless Sensor Networks) protocol [8] implements the same idea of probabilities for CH selection based on initial, remaining energy level of the nodes and average energy of network. In EDDEEC, use the concept of three levels heterogeneity and characterized the nodes as normal, advance and super nodes. EDDEEC is adaptive energy aware protocol which dynamically changes the probability of nodes to become a CH in a balanced and efficient way to distribute equal amount of energy between sensor nodes. At start of each round, node decides whether to become a CH or not based on threshold calculated by the following equation

$$T(s_i) = \begin{cases} \frac{p_i}{1-p_i(r \bmod \frac{1}{p})} & \text{if } s_i \in G \\ 0 & \text{otherwise} \end{cases}$$

where, G is the set of nodes eligible to become CH for round r and p is the desired percentage of CH.

$$p_i = \begin{cases} \frac{p_{opt} E_i(r)}{(1+m(a+m_o b))E(r)} & \text{if } s_i \text{ is the normal node} \\ \frac{p_{opt}(1+a)E_i(r)}{(1+m(a+m_o b))E(r)} & \text{if } s_i \text{ is the advanced node} \\ \frac{p_{opt}(1+b)E_i(r)}{(1+m(a+m_o b))E(r)} & \text{if } s_i \text{ is the super node} \end{cases}$$

which defines probabilities to become CH for current round. The expression is to distribute energy consumption over network efficiently, increase stability period and lifetime of network.

9) K-ways spectral clustering protocol in [9] it use the principle of spectral clustering method based on the notion of similarity matrix and using the eigenvectors differently. The k-ways approach is one of these methods and consists in dividing data into K disjoint classes based on the K eigenvectors related to K largest eigenvalues of a Laplacian matrix. The cluster head in each round of communication will be at a random position on the cluster. After so period nodes die at random locations of the network. Each round r , algorithm use the number $ck=(r \bmod |S_k|)$ to select the suitable cluster head for the appropriate cluster; where $|S_k|$ represents the total number of nodes in a defined cluster k . Besides, if the residual energy Eng_{ck} of the node, with $id=ck$, is greater than a threshold ΘEng , this node will be the CH of the cluster k in the round r . We define ΘEng as the minimum residual energy required for a given node to be a CH. Total consumed energy in each round.

$$E_{Round} = L(2N \cdot E_{elec} + N \cdot E_{DA} + K \cdot E_{mp} \cdot d_{toBS}^2 + (\sum_{i=1}^K d_i^4) \cdot E_{mp} + (\sum_{i=1}^{N-K-1} d_i^2) \cdot E_{fs})$$

We note that by considering K cluster, this energy depends on the distances between the CH and the non cluster heads of each cluster. i.e. E_{Round} is minimal if the quantity:

$$(\sum_{i=1}^K d_i^4) \cdot E_{mp} + (\sum_{i=1}^{N-K-1} d_i^2) \cdot E_{fs}$$

is minimal. However, it is known that the amplifier energy in a multipath fading channel models is greater than the amplifier one in a free space model (i.e. $E_{fs} \leq E_{mp}$). In addition, we have $d_j < d_0 \leq d_i$. Thus, to minimize the formula given in equation (12), all non CH nodes must operate in a free space model. The objective function that allows to decide whether to reconsider the partitioning process or not of the WSN, is defined by the distance matrix M_{dis} ($M_{dis} = [dis_{ij}]$; with dis_{ij} is the distance between the node i and the node j of the cluster labelled k (k) of each cluster. The allowed threshold to this function is d_0 . Hence, if at least one element of any M_{dis} is greater than d_0 , the considered number of clusters will be incremented ($K+1$) and the k-mean algorithm will be reused. Otherwise, the optimal number of clusters is K .

[10] ECHERP (Equalized Cluster Head Election Routing Protocol) proposed in [10] which pursues energy conservation through balanced clustering. Models the network as a linear system and using the Gaussian elimination algorithm, calculates the combinations of nodes that can be chosen as cluster heads in order to extend the network lifetime. nodes in the network is sent advertisement message to the BS, allowing the BS to have a global knowledge of the network. At this step, all the nodes are cluster head candidates, and each node has a unique ID. the BS runs the Gaussian elimination algorithm and computes the number of rounds at which

every node can be a cluster head. The BS broadcasts the unique IDs of the newly selected cluster heads, and their cluster members and the nodes use this information to form and enter a cluster. Each cluster head creates a TDMA schedule and broadcasts this schedule to the nodes in its cluster, in order to inform each node of the timeslot that it can transmit. When a CH change in a cluster position or in the total dissipation of a node residual energy, the BS uses again the Gaussian elimination algorithm to determine the appropriate cluster head election. **[11] EEHC (Energy Efficient Hierarchical Clustering)** was earlier proposed in Ref. [11] Protocol is designed with significant probabilistic clustering algorithm. The objective of this algorithm was to address the shortcomings of one-hop random selection algorithm in LEACH protocol by extending the cluster architecture to multiple hops. It is a distributed, k -hop hierarchical clustering algorithm to aiming at the maximization of the network lifetime. Initially, each sensor node is elected as a CH with probability " p " and announces its election to the neighboring nodes within its communication range. The above CHs are now called the "volunteer" CHs. Nodes present within " k "-hops distance from a "volunteer" CH, are supposed to receive the election message either directly or through intermediate forwarding. Next, any node that receives such CH election message and is not itself a CH, becomes a member of the closest cluster. In Addition, a number of 'forced' CHs are elected from nodes that are neither CHs nor belong to a cluster. Specifically, if the election messages do not reach a node within a preset time interval t , the node becomes a "forced" CH assuming that it is not within k hops of all volunteer CHs.

Table 1 shows the comparison of different protocols with some metrics related to the energy saving.

Table 1 Comparison of the protocols with some parameters

Protocol	Latency	Extra Overhead	Time Complexity	Node Mobility	CH Selection	CH Rotation	Multi level hop
LEACH(1)	Low	-	Stable	Limited	Probability/Random	Yes	No
M-LEACH(2)	Medium	Optimal Path Selection	Stable	Limited	Probability/Random	Yes	Yes
LEACH-C(3)	Low	GPS	Variable	Limited	Probability/Random	Yes	No
V-LEACH(4)	Low	Vice-CH maintenance	Variable	No	Probability/Random	No	No
TEEN(5)	Medium	Threshold implementation and attribute based naming	Variable	No	Random	Yes	Yes
APTEEN(6)	Medium	Threshold implementation and attribute based naming	Variable	No	Random	Yes	Yes
DECSA(7)	Medium	Formation of Base station CH and ordinary CH	Variable	No	Random	Yes	Yes
EDDEEC(8)	Low	Adv. Node and Super Node formation	Variable	Limited	Random	Yes	No
K-Way(9)	Low	GPS & CH formation	Variable	No	Random	Yes	No
ECHEP(10)	Medium	-	Variable	No	Random	No	Yes
EEHC(11)	Medium	-	Variable	No	Probability / Random	Yes	Yes

IV. CONCLUSION

Hierarchical routing protocols specially clustering techniques play a important role for saving energy in wireless sensor networks. In this survey we present the main characteristics of the most significant protocols that were proposed till now. And also hierarchical cluster structures facilitate the efficient data gathering and aggregation independent to the growth of the WSN, and generally reduce the total amount of communications as well as the energy spent. Clustering algorithms that adopt as primary election criteria other classical parameters like connectivity, distance, nodes' proximity, etc., have also been developed and relevant protocols are still being used, leading probably to more qualitative output such as well balanced clusters, etc. However the time complexity of these algorithms is difficult to be kept low as in leading probabilistic /random clustering algorithms, however, keeping the total overhead low, and reduce the latency the further development of efficient recovery protocols in case of CHs failure.

REFERENCES

- [1] W. R. Heinzelman, A .Chandrakasan, and H. Balakrishnan., *Energy-Efficient Communication Protocol for Wireless Microsensor Networks*. IEEE. Published in the Proceedings of the Hawaii International Conference on System Sciences, January 4-7
- [2] H. Zhou, Z. Jiang and M. Xiaoyan, "Study and design on cluster routing protocols of wireless sensor networks", 2006.

- [3] W.B. Heinzelman, A.P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE transactions on wireless communications*, Vol. 1, NO. 4, pp. 660-670, Oct. 2002.
- [4] M.B. Yassein, A. Al-zou'bi, Y. Khamayseh, and W. Mardini, "Improvement on LEACH protocol of wireless sensor network (VLEACH)", *International Journal of Digital Content Technology and its Applications*, Vol. 3, No. 2, pp. 132-136, Jun. 2009.
- [5] A. Manjeshwar and D.P. Agrawal, "TEEN: a protocol for enhanced efficiency in wireless sensor networks", in *Proceedings of the 1st International Workshop on Parallel and Distributed Computing Issues in Wireless Networks and Mobile Computing*, San Francisco, CA, Apr. 2001.
- [6] A. Manjeshwar and D.P. Agrawal, "APTEEN: a hybrid protocol for efficient routing and comprehensive information retrieval in wireless sensor networks", in *Proceedings of the 2nd International Workshop on Parallel and Distributed Computing Issues in Wireless Networks and Mobile computing*, Ft. Lauderdale, FL, Apr. 2002.
- [7] Yonga, Z. and Peia, Q., "A Energy Efficient Clustering Routing Algorithm Based on Distance and Residual Energy for Wireless Sensor Networks", *International Workshop on Information and Electronics Engineering (IWIEE)*, Harbin, China, March 10 -11, 2012.
- [8] Nadeem Javaid, T. N. Qureshi, A. H. Khan, A. Iqbal, E. Akhtar, M. Ishfaq: EDDEEC: Enhanced Developed Distributed Energy-efficient Clustering for Heterogeneous Wireless Sensor Networks. *ANT/SEIT 2013*: 914-919
- [9] Ali Jorio, Sanaa El Fkihi, Brahim Elbhiri, Driss Aboutajdine , "A New Clustering Algorithm in WSN Based on Spectral Clustering and Residual Energy", *The Seventh International Conference on Sensor Technologies and Applications*, August 25-31, 2013. pages – 119-125
- [10] Stefanos A. Nikolidakis, Dionisis Kandris, Dimitrio, D. Vergados, and Christos Douligeris, "Energy Efficient Routing in Wireless Sensor Networks Through Balanced Clustering," www.mdpi.com/journal/algorithm (2013)
- [11] S. Bandyopadhyay and E. Coyle, An energy efficient hierarchical clustering algorithm for wireless sensor networks, in *22nd Annual Joint Conf. of the IEEE Computer and Communications Societies (INFOCOM 2003)*, San Francisco, CA, April 2003.