# A SURVEY ON CLUSTER BASED ENERGY EFFICIENT ROUTING PROTOCOLS IN WIRELESS SENSOR NETWORK

[1]Binkal Ahir, [2]Bintu Kadhiwala
[1]P.G. Student, Co-Department Scet Surat.
[2] Assistant Professor, Co-Department Scet Surat.
Email:[1]binkalahir3@gmail.com, [2]bintu.kadhiwala@scet.ac.in

**Abstract— In recent times, there has been a growing interest in Wireless sensor networks (WSNs). They become gradually attractive in wide variety of applications due to their low cost, small in size, low power, and self-organizing behavior in harsh environment. The major problem of WSNs research is Routing protocol. Data aggregation reduces the number of message to be transmitted in the network and provide fused information to the Base Station (BS). Based on network structure, routing protocols that support data aggregation can be divided into two categories: flat routing and hierarchical or clustering routing. Due to a variety of advantages, clustering techniques are becoming an active branch of routing technology in WSNs. Hierarchical clustering algorithms are very important in increasing the network's life time. In this paper, we present a general classification of published clustering protocol. We survey different existing clustering algorithms for WSNs; highlighting their objectives, main feature, issues, etc. We also compare these clustering algorithms based on metrics such as clustering method, network classification, Cluster head (CH) selection criteria, multi path support, load balancing, scalability and energy efficiency.**

**Index Terms—Cluster construction, Cluster heads, Data aggregation, Clustering protocols.**

## I. INTRODUCTION

WSNs are consisting of light weight and tiny sensor nodes with limited power, communication, storage, and computation capabilities. Sensors are capable of detecting physical constraints such as temperature, sound, pressure and humidity. The sensor nodes are deployed in the sensing area through wireless links which provide opportunities for many civilian and military applications, for example: intrusion detection, availability of equipments, environment observation, home intelligence, biomedical sector, building monitoring and target tracking. Even monitoring the health status of cattle stocks on farms is supported by WSNs. In disaster management situations such as earthquakes, sensor networks can be used to selectively map the affected regions [1].

As described in [2], a sensor networks is classified in to two types on the basis of their mode of functioning and the type of target application.

• *Proactive Networks*: In this network, the sensor nodes periodically switch on their transmitters, sense the attribute of environment and transmit the data of interest. Thus, they are well suited for applications requiring periodic data monitoring.

• *Reactive Networks:* In this scheme the nodes react immediately to sudden and drastic changes in the value of a sensed parameter. As such, they are well suited for time critical applications.

Once sensor nodes are deployed, they form a network through short-range wireless communication. One central application of

WSNs is data gathering, i.e., sensor nodes transmit data, possibly in a single hop or multi-hop fashion, to the BS. Actually, one is often only interested in collecting a relevant function of the sensor measurements at the BS, rather than taking the data from all the sensors. Often, sensors are deployed in an environment where data generated by neighboring sensors are highly co-related or even redundant due to high density in the network topology. Furthermore, the nature of the physical phenomenon constitutes the temporal correlation between each successive reading of a sensor node. Hence, it is necessary to define the capacity for combining this data in to high quality information at the sensors or intermediate nodes, and transporting only specific functions of sensor measurements to the BS resulting in preservation of energy and bandwidth. Since in-network aggregation plays a key role in improving such capacity for WSNs, we can reasonably call such capacity as aggregation capacity for WSNs [3].

Based on network structure, routing protocols that support data aggregation can be divided into two categories: flat routing and hierarchical routing. In a flat topology, data transmission is performed hop by hop usually in the form of flooding. The example of flat routings in WSNs is Flooding and Gossiping, Direct Diffusion. Flat routing protocols are relatively effective for small-scale networks. On the other hand, in a hierarchical topology, nodes are organized into lots of clusters according to specific requirements. Generally, each cluster comprises a leader called CH and other nodes are member nodes (MNs). CH acts as a gateway for collection and forwarding the aggregated packets to the BS. Other nodes act as MNs and perform the task of information sensing. Based on the functionality and selection of aggregator, there are four Data aggregation techniques in hierarchical routing [4].

*A. Chain based data Aggregation*

In chain based data aggregation the data is sent only to the closest neighbor. The cluster formation and CHs selection are not performed. Each node must know the location of all other nodes in the network. Each node determines the distance to its neighbors using the signal strength and then adjusts it to communicate only with the closest neighbor. Collected data moves across the nodes, gets aggregated at each node, and eventually, a single chosen node transmits data to the base station. Nodes take turns in transmitting to the BS so that the power dissipation for communicating with the BS is uniformly distributed among all the nodes. The chain construction is done in greedy fashion with the assumption that all the nodes have global knowledge of the network [5].

*B. Grid based Data Aggregation*

In this approach, the data forwarding can be done in three processes the cluster grid construction process, query forwarding and data forwarding. An aggregator is selected based on the geographical position with respect to either sink or grid center. The aggregator is fixed for grid and it aggregates the data from all the sensors. Hence, the sensors within a grid do not communicate with each other [5].

*C. Tree based Data Aggregation*

The path is computed centrally using BS or it can be computed by running shortest path algorithm at particular node. The path information is broadcasted to the network. In collection phases, all the leaf nodes forward data to its parent and then it roots towards the sink. Any node failure in the root blocks data and increases latency with the decrease in packet delivery ratio [5].

*D. Cluster based data Aggregation*

Cluster based data aggregation technique, solve the problem such as transmission delay and loss of data caused due to node failure in the root to sink. In large sized network, there is a need to find optimal path for the efficient communication of data to the sink. It increases communication cost and reduces the efficiency. In such case, instead of communicating data individually to sink, it can be aggregated at cluster head node, and that compressed data is transmitted to sink [5].

In this paper, we summarize different clustering protocols. The paper is organized as follows: The section II describes clustering process in sensor network. In section III, we present various clustering protocols and their comparisons followed by conclusion.

## II. CLUSTERING IN SENSOR NETWORK

Clustering routing is becoming a popular branch of routing technology in WSNs due to variety of advantages, such as high scalability, data aggregation/fusion, less load, more robustness, less energy consumption.

### A. Set Up Phase

Set Up phase consist other 2 stages. It starts by the CHs selection stage and proceeds by constructing clusters.

1. CH selection

The first step in the set up phase is the selection of CHs. CHs are responsible for the collection of information within the cluster, data aggregation process and transmission of fused information towards next hop or BS. So, CHs selection plays a significant role. CHs are selected using three methods.

a. Distributed
b. Centralized
c. Hybrid

In the Distributed methods, inter-CHs coordination is performed in a distributed manner and each individual CH takes charge of forming its own cluster. Either stochastically or involving some probabilistic resource parameters, nodes are participating to become CH. There are several advantages of distributed cluster-based algorithms, but since a single node does not have knowledge of the topology and characteristics of the entire network, distributed methods provide no guarantee about the fair placement of CHs and optimum number of CHs are not guaranteed. There is lot of overhead at node due to transmission of large number of control messages on restricted resources of sensor nodes, which reduce the overall network lifetime [ 4, 5, 6, 7,8].

In the Centralized method, the CHs are selected by the central authority like BS. This method provide fair placement of the optimum number of CHs and to mitigate energy expenditure of nodes in re-clustering stages of a balanced cluster. The burden of CHs selection and cluster formation phase are taken by BS who has the unlimited energy source and high processing capabilities. However, this requires periodically each sensor node in the network to send necessary information like remaining energy level to the central authority [4,5,8,9].

Hybrid method use best feature of above two methods. This method is found especially when CHs are rich in resources [4].

2. Cluster Formation

Once the CHs are selected, next step is cluster formation. The selected CHs broadcast the advertisement message to other nodes. Each node receives this message and send join message to the nearest CH node. The main functionality of this step is to manipulate the size of the clusters, to minimize and balance the energy expenditure in the network, to detect faults and recover from failing situations or in the event-driven clustering schemes, trigger the cluster formation stage only when and where it is needed [4,5,6,7,8,9].

### B. Data Transmission Phase

The setup phase is followed by the steady data transmission phase, which starts with data aggregation at CHs and data transmission stages.

1. Data Aggregation

The data generated by sensor node are highly redundant and co related with each other due to spatial and temporal co-relation. However, this redundancy entails generation of large numbers of highly analogous data, which imposes high level of energy expenditure into the network to be processed and forwarded to the BS. Therefore, to save limited energy of WSNs, Data aggregation mechanisms are proposed. The main objective of data aggregation is to eliminate redundant transmission of data and provide fused information to the BS in order to increase the network lifetime. However, data aggregation may degrade some Quality characteristics of the network such as data accuracy and latency. To achieve the optimal trade-off, data aggregation techniques should be closely coupled with data routing protocols in order to have complete domination on different forwarding paradigms to promote in-network data aggregation capacity. The CHs perform simple aggregation functions like MIN, MAX, SUM, AVG and XOR to fuse data [3,4,5,6,7,8].

2. Data Transmission

In this phase, as coordinators of the cluster, the CHs transmit the aggregated data to the BS for further analysis by the end user according to the type of the application. In single-hop transmission, the transmission of a packet from sensor nodes to the CH and from CHs to the BS can be done through direct transmission. In multi-hop transmission, the CHs transmit data to higher level node or by assistance of other nodes in the path. The amount of energy used for transmission is depends upon the distance between source to destination and packet size. Radio Model is used for energy calculation [4,5,6,7,8].

## III. CLUSTERING PROTOCOLS

There are several cluster based protocols in the literature. Some of them are described here.

### *A. Low Energy Adaptive Clustering Hierarchy (LEACH)*

LEACH is proposed by Heinzelman, Anantha P. Chandrakasan and Hari Balakrishnan [6]. It is one of the popular clustering routing algorithm for WSNs. The operation of LEACH is broken up into lots of rounds; during the set-up phase, each node decides whether to become a CH or not for the current round. This decision is based on the number of times the node has been a CH so far and the suggested percentage of CHs for the network. This decision is made by the node choosing a random number between 0 and 1. Then, node becomes a CH for the current round if the number is less than the threshold formula shown in (1).

$$T(n) = \begin{cases} \dfrac{P}{1 - P\left(r \bmod \frac{1}{P}\right)}, & \text{if } n \in G \\ 0, & \text{otherwise} \end{cases}$$

(1)

Where P is the desired percentage of CHs, r is the current round and G is the set of nodes that have not been elected as CHs in the last 1/P rounds. CHs rotation is performed at each round. During the steady-state phase, the sensor nodes sense the environment attribute and transmit data to the CHs. The CHs aggregate data arriving from nodes that belong to the respective cluster, and send an fused or aggregated packet to the BS directly. LEACH uses a TDMA/code-division multiple access (CDMA) MAC to reduce inter-cluster and intra-cluster collisions. After a certain time period, which is determined a priori, the network goes back into the set-up phase again and enters another round of CHs election.

The advantages of LEACH include: (1) Any node that served as a CH in certain round cannot be selected as the CH again, so each node can evenly share the load imposed upon CHs to some extent; (2) Use of a TDMA and CDMA schedules prevents CHs from unnecessary collisions.

However, there exist a few disadvantages in LEACH are: (1) The CHs are directly communicated with BS so performs the single-hop inter-cluster, which is not applicable to large-region networks. Long-range communications directly from CHs to the BS can cause too much energy consumption; (2) CHs are elected in terms of probabilities without energy considerations; (3) Since CH election is performed in terms of probabilities, CHs are not distributed uniformly throughout the network.

### *B. Energy Efficient and Balanced Cluster-Based Data Aggregation Algorithm for Wieless Sensor Networks (EEBCDA)*

EEBCDA is proposed by Jun Yue, Weiming Zhang and Weidong Xiao [7]. This protocol solves the problem of Un balanced energy dissipation in cluster based and homogeneous WSNs. The cluster head transmit data to Base station in one hop communication. It divides the network into rectangular grids with unequal size and makes cluster head rotate among the node in each grid respectively. The CHs in the grid which are further away from BS consume more energy in each round, these grids have more nodes to participate in CHs rotation and share energy load, so this protocol is able to balance energy dissipation on a long view. The CHs are selected in distributed manner. Initially each node send node id, location in grid and Energy level to other nodes in the grid. The higher energy nodes are selected as CHs. For the sake of CHs selection in next round, each member transmits its residual energy along with its data to CH at the last time of data gathering in every round.

The advantage of this protocol is that it balance the energy consumption by an energy efficient way and prolong network life time.

The disadvantage of this protocol include: (1) The CHs transmit data to BS by one hop communication. So this is not efficient in large area network as CHs are inefficient to transmit to BS directly; (2) There is problem of unbalanced energy dissipation if CHs are communicated with multipath. The setup phase is followed by the steady data transmission phase, which starts with data aggregation at CHs and data transmission stages.

There are several cluster based protocols in the literature. Some of them are described here.

### *C. Grouping of Clusters for Efficient Data Aggregation in wireless sensor network (GCEDA)*

GCEDA is proposed by Dnyaneshwar Mantri, Neeli R Prasad and Ramjee Prasad [8]. The GCEDA algorithm operates in three phase. Cluster formation, intra-cluster and inter-cluster aggregation with grouping of nodes and CHs for communication of aggregated data packets to sink.

In this protocol the group based data aggregation method is proposed, where grouping of nodes based on available data and correlation in the intra-cluster and further grouping of CHs at the network level help to reduce the energy consumption. MNs transmit data to CHs and CHs again perform aggregation at higher level and transmits aggregated information to sink. While transferring data to sink, it considers multi-hop communication and CHs groups according to information of aggregated data packets. The nodes are uniformly distributed and it select the CHs based on highest energy, minimum distance to sink calculated using Euclidean distance and the highest number of neighbor nodes. CHs groups according to available data from each CHs to perform the further aggregation for communicating to sink. Grouping of nodes in intra-cluster and grouping of CHs at inter cluster reduces the data packet count at the sink. It reduces the effective energy required, which prolongs the network lifetime.

The advantage of GCEDA is: (1) Uniformly distributed node in each cluster so each cluster is balanced; (2) Inter cluster aggregation is also performed.

### *D. Centralized Energy Efficient Clustering (CEEC)*

CEEC is proposed by M. Aslam, N.Javid and A.Rahim [9]. In CEEC whole network area is divided into three equal regions, in which nodes with equal energy are spread in same region. The network model contains three different types of nodes called normal, advance and super nodes. These nodes preserve different levels of energy. As the distance of nodes from BS increases, energy level of the nodes is also increases. It brings equal distribution of resources with respect to responsibilities of nodes. The differentiate feature of this model is that nodes associate with their own type of cluster head nodes.

BS centrally selects optimum number of Cluster Heads. The CHs are selected based on four parameter initial energy of node, residual energy of nodes, average energy of each region and location of nodes. After completion of one round each node send these four parameters to BS. Operation of CEEC is based on rounds, with adjustable duration. Each round is divided into Network Setting Time (NST) and Network Transmission Time (NSS). During NST, CHs are selected and multiple clusters are formed. During NTT, sensed information from all nodes is transmitted to BS with help of CHs.

The advantage of this protocol is it guarantees the optimum number of CHs in each round as they are selected by BS.

The disadvantage of this protocol include: (1) The CHs are directly communicated with the BS. This is not suitable for large homogeneous network; (2) Start Up Energy dissipation is more.

### *E. Threshold sensitive Energy Efficient sensor Network protocol (TEEN)*

TEEN is proposed by Arati Manjeshwar and Dharma P. Agrawal, is a hierarchical protocol [2]. The main goal is to cope with unexpected changes in the sensed attributes as like temperature. The nodes sense their environment continuously, but it transmit whenever it is required so the energy consumption in this algorithm can potentially be much less than that in the proactive network, due to less data transmission.

In TEEN, CHs are selected as like in LEACH protocol, a 2-tier clustering topology is built and two thresholds, hard threshold (HT) and soft threshold (ST), are defined. The HT is a threshold value for the sensed attribute, is the absolute value of the attribute beyond which, the node sense the value must switch on its transmitter and report the sense data to its CH. The ST is a small change in the value of the sensed attribute which triggers the node to switch on its transmitter and transmit.

TEEN has the following advantages: (1) Only the sensitive data we demand can be transmitted, so that it reduces the unnecessary energy transmission consumption and improves the effectiveness of the receiving data; (2) TEEN is match for reacting to large changes in the sensed attributes, which is suitable for reactive scenes and time critical applications.

However, there exist a few drawbacks in TEEN are as follows: (1) It is not suitable for periodic reports applications since the user may not get any data at all if the values of the attributes may not reach the threshold; (2) There exist wasted time-slots and a possibility that the BS may not be able to distinguish dead nodes from alive ones; (3) There is complexity while constructing cluster at higher levels.

### *F. Adaptive Periodic Threshold-sensitive Energy Efficient Sensor Network Protocol (APTEEN)*

APTEEN introduced by Arati Manjeshwar and Dharma P. Agrawal, is an extension to TEEN and aims at both transmitting periodic data as well as reacting to time critical events [10]. It is Hybrid protocol provide both functionality of proactive and reactive networks. The protocol adjust the parameters issued by the cluster head, parameters can be changed according to the needs of the user ,like HT, ST, Counting time (CT) is the most time period represented successful data transmission of a node. If a node does not send data for a time period equal to the CT, it must sense and transmit the data again.

The advantages of APTEEN include: (1) APTEEN combines both proactive policies, which is similar that of LEACH, and reactive policies, which is similar that of TEEN. Accordingly it is suitable in both proactive and reactive applications; (2) It provide flexibility by setting the count-time interval, and the threshold values for the energy consumption can be adjusted by changing the count time as well as the threshold values.

The main disadvantages of APTEEN are as follows: (1) There exist additional complexity required to implement the threshold functions and the count time; (2) Both TEEN and APTEEN share the same drawbacks of additional overhead and complexity of cluster construction in multiple levels, implementing threshold-based functions, and dealing with attribute-based naming of queries-APTEEN more than TEEN.

### *G. Well Balanced Threshold sensitive Energy Efficient sensor Network protocol (WB-TEEN)*

Hierarchical clustering algorithm WB TEEN and WBM-TEEN (WB-TEEN with Multihop Intracluster) are proposed by Zibouda Aliouat and Saad Harous [12]. Each cluster has almost equally number of member node so cluster is balanced and the total energy consumption between sensor nodes and cluster heads is minimized by using multihop intra cluster aggregation. The CHs are periodically elected depending on their residual energy level.

The protocol WB-TEEN is an improvement of protocol TEEN which enables clusters balancing means it avoids clusters formation with a significant difference in size. The CHs calculates its degree using formula (3) and according to this number it accepts the membership request of other nodes.

$$Degree = [(NN - CH_{nbr})/CH_{nbr}] + 1$$

(2)

Where CHnbr is a number of CHs and NN is a total number of nodes in the network.

The advantage of this protocol is that it combines the best feature of the LEACH and TEEN protocol. In the cluster nodes are evenly distributed so clusters are energy balanced.

The disadvantage of this protocol is that there is overhead to select the CHs. If the node is close to one CH and if that CH has enough members then also the node has to join to the other cluster which may be far away from it.

The Table 1 shows the comparison of different clustering protocols in terms of different key parameter such as clustering method as how CH are selected is important, network classification whether it is proactive or reactive protocol, initial energy of all node is same or different also play crucial role in life time of the network , various CH selection parameters as being CH many tasks has to be performed, scalability, energy load balancing of each cluster, CH are communicate to BS via single hop or multi hop and energy efficiency. Both Centralized and Distributed methods have their own advantages and disadvantages. However, there is need of clustering algorithm that use the best feature of both method and provide good network lifetime.

**Table I. COMPARISION OF DIFFERENT CLUSTERING PROTOCOLS**

| Protocol Name | Clustering Method | Classification | Initial Energy of node | CH Selection criteria | Scalability | Load Balancing | Multi Hop | Energy Efficiency |
|---|---|---|---|---|---|---|---|---|
| LEACH | Distributed | Proactive | Homogeneous | Initial Energy | Very Low | Low | No | Very Low |
| EEBCDA | Distributed | Proactive | Homogeneous | Initial Energy, Residual Energy | Medium | High | No | Medium |
| CEEC | Centralized | Proactive | Heterogeneous | Initial Energy, Residual Energy, Location, Average Energy of network | Medium | Medium | No | Medium |
| GCEDA | Centralized | Proactive | Homogeneous | Residual Energy, Location, No. of Neighbors | Medium | Medium | Yes | Medium |
| TEEN | Distributed | Reactive | Homogeneous | Initial Energy | Low | Medium | Yes | High |
| APTEEN | Centralized | Hybrid | Homogeneous | Residual Energy, Location | Medium | Medium | Yes | Medium |
| WB TEEN | Distributed | Reactive | Homogeneous | Initial Energy | Medium | High | No | High |

## IV. CONCLUSION

WSNs have fascinated significant attention over the past few years. In this paper, we classified the different clustering protocols. We summarized recent clustering algorithm, stating their strength and limitations. We hope that this will support protocol researchers to take into account the various characteristics of the clustering routing methods while designing an energy efficient Clustering routing protocol.

REFERENCES


[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, “Wireless sensor networks: a survey,” ELSEVIER, 2001.

[2] A. Manjeshwar and D.P. Agrawal, “TEEN: A Routing Protocol for Enhanced Efficiency in Wireless Sensor Networks,” IEEE , 2001.

[3] C. Wang, C. Jiang, Y. Liu, X.Y. Li, S. Tang, H. Ma, “Aggregation Capacity of Wireless Sensor Networks:Extended Network Case,” IEEE INFOCOM , 2011.

[4] A.A. Abbasi and M. Younis, “A survey on clustering algorithms for wireless sensor networks,” ELSEVIER, 2007

[5] R. Rajagopalan, P.K. Varshney, “Data-aggregation techniques in sensor networks: a survey,” Communications Surveys & Tutorials, IEEE , vol.8, no.4, pp.48-63, Fourth Quarter 2006.

[6] 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, October 2002

[7] J. Yue, W. Zhang, W. Xiao and D. Tang, “Energy Efficient and Balanced Cluster-Based Data Aggregation Algorithm for Wireless Sensor Networks,” ELSEVIER International Workshop on Information and Electronics Engineering(IWIEE), 2012

[8] D. Mantri, N. R. Prasad and R. Prasad, “Grouping of Clusters for Efficient Data Aggregation (GCEDA) in Wireless Sensor Network,” IEEE International Advance Computing Conference (IACC) , 2013

[9] M. Aslam, T. Shah and N. Javaid, “CEEC: Centralized Energy Efficient Clustering A New Routing Protocol for WSNs,” IEEE, 2013

[10] A. Manjeshwar and D. P. Agrawal, “APTEEN: A Hybrid Protocol for Efficient Routing and Comprehensive Information Retrieval in Wireless Sensor Networks,” Proceedings of the International Parallel and Distributed Processing Symposium (IPDPSí02) IEEE, 2002

[11] Z. Aliouat and S. Harous, “An Efficient Clustering Protocol Increasing Wireless Sensor Networks Life Time,” IEEE International Conference on Innovations in Information Technology (IIT) , 2012