A Survey: Cluster based Routing Protocols in Wireless sensor network

[1]Rohan Jaysinh Parmar, [2]Dr. Sanjay kumar
[1]P.G. Student, Jaipur National University, Jaipur.
[2] Associate Professor, Jaipur National University, Jaipur.

Abstract— Due to wide range of applications, low cost, small in size, low power, and self-organizing behavior, Wireless sensor network is a popular subject in research community. Wireless sensor network node has low energy, due to this, energy consumption of node during data aggregation plays major role in Wireless sensor network. Routing protocol with large overhead of data aggregation is not feasible for Wireless sensor network. Routing protocol with low energy consumption in sensor network is, therefore, a challenging task. Clustering mechanism for routing has many advantages over hierarchical routing, therefor, clustering techniques are becoming an active branch of routing technology in WSNs. This paper describe the present situation of the art in data aggregation using clustering technique for WSNs. Various type of routing protocol which used clustering techniques for data aggregation are discussed and their comparison and problem based on parameter such as clustering technique, network classification, process of Cluster head (CH), support of multi path, load balancing, scalability and energy efficiency are presented. Analysis on future research work for routing protocol is also included.

***Index Terms* — WSNs, Cluster formation, Cluster heads, Data aggregation, Protocols.**

I. INTRODUCTION

Collection of hundreds or thousands of sensor node is known as sensor network[1], which is communicate to gather to achieve particular task. Advance research in Micro-Electro-Mechanical Systems (MEMS) provides the mechanism for development of smart sensors. Sensor nodes are tiny in size, have limited processing and computing power, very low communication and storage capabilities. We have variety of sensors which are capable to measure properties of the environment. One or more sensors are attached with WSNs node. WSNs nodes are deployed in the remote area through wireless links. Today many applications are use WSNs for security, monitoring and other purpose [2].

According to their operating mechanism and targeting application, sensor networks are classified in to two types [3].

(1) *Proactive Networks*: It is fishable for the applications which require monitoring the environment at regular interval because in this type of network, the sensor nodes, at regular interval, sense the environment of interest, switch on the transmitter and transmit the data of interest.

(2) *Reactive Networks:* It is fishable for the application which is time critical because in this type of sensor network, nodes react immediately to sudden and drastic changes in the environment of interest.

For creation of wireless sensor network, we first deploy the nodes as required by application and nodes form a network using short–range wireless communication. Generally WSNs are used in a harsh environment and so generally we can't have deployment control over it. Given large area to be covered, the small battery life, and possibility of having damage during deployment, hundreds or thousands of sensors are expected in WSNs. Most of applications of WSNs is attribute gathering, i.e. sensor node sense environment and transmit required attribute in single or multi-hope manner to Base station. Most of applications are interested in correlated data of the sensor measurements at BS instead of collecting attributes from all nodes. As we stated above there is high density of nodes are present in WSNs, it is likely that information are co-related or redundant. Due to this type of data, in a WSNs data aggregation plays a major role. By data aggregation one can combine the data at intermediate node and can present high quality of information. Such kinds of capability of WSNs are known as aggregation capacity for WSNs [4].

We can categorize routing protocols that support data aggregation in two categories Based on network structure. (1) Hierarchical routing: In this topology, based on requirement, nodes form a cluster. Each cluster has CH which act as leader and that CH perform the collection and aggregating the data packet and forward that packet to BS. (2) Flat routing: In this topology data transmission is done hope by hope, in the form of flooding. Flooding and Gossiping, Direct Diffusion is examples of a flat routing. If we have small scale network than flat routing is fishable. We are interested in a hierarchical routing so we will discuss about it throughout this paper. We can categorize hierarchical routing in four categories [5].

A. Chain based data Aggregation [12]

In this method energy efficient chain is created for data

ISSN: 2278 – 1323

transmission. Different researchers have proposed different algorithm for chain based data aggregation. For chain based data aggregation algorithm, all nodes in network must know the location of other nodes in the network. In this technique data is transmitted to its near neighbor. To determine the nearest neighbor, signal strength is used. Collected data travels on network, aggregated at each node and only chosen nodes transmit data to a base station. All nodes periodically get turn on for transmit the data to base station so that power consumption for transmitting the data is distributed among the all nodes.

B. Grid based Data Aggregation

In grid based data aggregation, network is divided in the grid and each grid has fixed aggregator node and that node performs the data aggregation. Based on the geographical position, an aggregator is selected. In this type of network node does not communicated with each other, only communicate whit aggregator node [6].

C. Tree based Data Aggregation

In tree based data aggregation, tree is form to aggregate the data. This type of aggregation technique is used when we want combine data from all nodes like radiation effected area and other application. In this concept, tree is formed and data aggregation is done at each intermediate note and aggregated data is transfer toward the root node [6].

D. Cluster based data Aggregation

In Cluster based data aggregation technique, group of node form a cluster and one leader node, known as cluster head, aggregate the data and transmit it to base station. Usage of this method reduces the amount of data transmitted to base station and so it increase the life time of network. Clustering routing protocol become popular due to its grate advantage like data aggregation/fusion, robust, more scalability, energy friendly etc.

In this paper, we analyze different clustering protocols. The paper is structured as follows: The section II gives the detail of clustering process. In section III, we describe the different clustering protocols their analysis and comparisons. In section IV we present the conclusion.

II. CLUSTERING PROCESS IN WSNs

All clustering routing protocol follow the step to form a cluster and data transmission. The process of cluster formation and data transmission are as follow:

A. Set Up Phase

Cluster set up phase is divided in to two part. First cluster head selection process is done and after that cluster formation is done.

I. CH selection

The initial step in set up phase is choosing the cluster heads. Gathering information from cluster member nodes, data aggregation, transmission of data to BS or next hope are duties of CHs. It means CHs selection plays important role. Following are methods which is used to select the CHs:

a. Centralized

b. Distributed

c. Hybrid

Each CH is responsible for forming its own cluster in the Distributed methods. Inter-CHs coordination is performed in a distributed manner for cluster formation. Some random or probabilistic parameter are used to select the CH. This technique has many advantages but all node does not know the network topology and other useful information of network and due to this technique does not provides guarantee of best CH placement and optimum number of Cluster head. In this technique large overhead are present which reduce the overall network lifetime. [5, 6, 7,8,9].

Centralized authority like BS select the CH in centralized CH selection technique. In this method CH selection is taken by centerline authority and as we know this authority has unlimited energy and processing power, burden of CH selection and cluster formation is done easily. By using this method balanced cluster and optimum number of Ch can be choose. To use this method, some useful information must be send to base station by each node like node location and remaining energy level. [5,6,9,10].

When CHs have rich resources, Hybrid method is used. This method take best things of both centralized and distributed method [5].

2. Cluster Formation

Second step is Cluster formation. After selecting the CHs, we form clusters. All CHs sends invitation massage to other node for join as cluster member. As soon as node receives the invitation massage, node respond on it and send joining massage to its near CHs. 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 [5,6,7,8,9].

B. Data Transmission Phase

After cluster formation step, next step is data transmission phase. This step starts with data aggregation and followed by data transmission.

1. Data Aggregation

Sensor node sense the environment and generate the data. The data generated by sensor node are highly redundant and co related with each other. If we forward all data to base station then it lead to high energy consumption because of large amount of data. If we use data aggregation techniques then we can save the energy of a WSNs. The main objective of data aggregation is to

ISSN: 2278 – 1323

eliminate redundant transmission of data and provide fused information to the BS in order to increase the network lifetime. Data aggregation may lead to data lost. One should use data aggregation technique according to application so that any loss not occurs and also energy can be saved. 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 data transmission phase, data are send to the base station by CHs for further analysis. End user analyses the data according to application requirement. 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. Energy consumption is dependent upon the amount of data and the distance between transmitter and receiver. 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 most popular clustering routing algorithm for WSNs. Operation of leach is divided in many rounds; In startup phase, all node take decision about become a CHs or not. The decision is based on 2 parameter. 1 how many times node become cluster head before this round 2 how many number of CHs will allow for this round. This decision is made by the node choosing a random number between 0 and 1. Then, if the number is less than the threshold then node become CHs the current round. Threshold formula shown in (1).

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

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. All cluster members sends data to its respective CHs and respective CHs aggregate data arriving from nodes, CHs sends aggregated data to the BS. To avoid inter-cluster and intra-cluster collisions, LEACH uses a TDMA/code-division multiple access (CDMA) MAC.

Main advantages of LEACH are: (1) Any node, Once served as a CH in certain round cannot be selected as the CH again, By this, lode imposed on CHs are shared by all nodes and which lead to long network life time (2) TDMS and schedules are used which stops the unnecessary collisions.

Disadvantages in LEACH are: (1) Single-hop inter-cluster communication is used by CHs to transfer the data, for large network this technique is not fishable; Long range communication between BS and CHs lead to more energy consumption (2) For CHs selection only probabilities consider without energy considerations; (3) CHs are not distributed uniformly throughout the network because CHs are chooses based on probability.

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

EEBCDA is proposed by Jun Yue, Weiming Zhang and Weidong Xiao [8]. In this protocol, unbalanced energy dissipation problem in cluster based and homogeneous WSNs is solved. BS and CHs uses one hope communication for data and massage transfer. Network is divided in grid with unique size and one grid act as one cluster and for life time CH is choose from particular grid only. 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 [9]. 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

ISSN: 2278 – 1323

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 [10]. 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, and 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,

ISSN: 2278 – 1323

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.

$$Degree = [(NN - CH_{nbr})/CH_{nbr}] + 1 \quad (2)$$

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
WB TEEN	Distributed	Reactive	Homogeneous	Initial Energy	Medium	High	No	High
TEEN	Distributed	Reactive	Homogeneous	Initial Energy	Low	Medium	Yes	High
CEEC	Centralized	Proactive	Heterogeneous	Initial Energy, Residual Energy, Location, Average Energy of network	Medium	Medium	No	Medium
APTEEN	Centralized	Hybrid	Homogeneous	Residual Energy, Location	Medium	Medium	Yes	Medium
GCEDA	Centralized	Proactive	Homogeneous	Residual Energy, Location, No. of Neighbors	Medium	Medium	Yes	Medium
EEBCDA	Distributed	Proactive	Homogeneous	Initial Energy, Residual Energy	Medium	High	No	Medium

IV. CONCLUSION

Many researchers get attention on WSNs on the past few years. In this paper, we focused and study the different clustering protocols. We provide the comparison and advantage and disadvantages of different algorithm. This study will be helpful for researcher to introduce new protocol.

REFERENCES

[1] J.M. Kahn, R.H. Katz, and K.S.J. Pister, "Mobile Networking for Smart Dust," Proc. ACM/IEEE Intl. Conf. Mobile Computing and Networking (MobiCom '99), Aug. 1999, also available at http:// robotics.eecs. erkeley.edu/pister/SmartDust/.

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

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

[4] 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.

ISSN: 2278 – 1323

[5] A.A. Abbasi and M. Younis, "A survey on clustering algorithms for wireless sensor networks," ELSEVIER, 2007
[6] 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.
[7] 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
[8] 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
[9] 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
[10] M. Aslam, T. Shah and N. Javaid, "CEEC: Centralized Energy Efficient Clustering A New Routing Protocol for WSNs," IEEE, 2013
[11] 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
[12] 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

ISSN: 2278 – 1323