ISSN No. 0976-5697

Volume 7, No. 3, May-June 2016

International Journal of Advanced Research in Computer Science

REVIEW ARTICLE

Available Online at www.ijarcs.info

Clustering in Wireless Sensor Networks: A Review

Shilpa Mahajan
Department of computer Science & Engineering
SHIVA Institute of Engineering & Technology
Bilaspur,India

Pushpender Kumar Dhiman
Department of Computer Science & Engineering
National Institute of Technology,
Hamirpur,India

Abstract: WSN is a collection of very tiny nodes which has gained a high attention of the researchers due to its various applications in different sectors. The tiny devices, called the sensor nodes are deployed in a harsh area and left unattended to continuously report the parameters of the environment, based on the application. The tiny nodes are constrained by the energy and hence it becomes necessary to consume less energy by using the means of clustering so that the lifetime of the whole network can be prolonged. This paper will represent an introduction to the clustering and various issues regarding to it.

Keywords: WSN, Sensor Node, Energy Efficiency, Clustering Algorithm.

1. INTRODUCTION

With advancement in the technology of MEMS (Micro Electro Mechanical System), the wireless communication and also the WSN are emerged. WSNs have become the most tremendous area of research. WSN is a collection of huge amount of tiny devices i.e. sensor nodes, distributed over a large area. Node is the smallest device of the WSN which is collaborated with the processing, communication and sensing capabilities. The basic components of the node are battery (limited), processor (limited), energy (limited), transceiver and small amount of memory. A node the WSN can be: (1) Homogeneous, and (2) Heterogeneous. **Homogeneous Node:** This is the type of node which is equipped with the same amount of communication, sensing and processing capabilities. **Heterogeneous Node:** This is the classification of node with variable sensing, communication and processing capabilities.

WSN are useful for monitoring various applications. Like surveillance, security, disaster management, military, healthcare and environmental studies. The WSN is a collaboration of the large number of nodes with limited energy, processing and communication capabilities which are deployed to perform an ad hoc operation without having any infrastructure and any central control point. The sensor node collects data locally from the target domains and then forwards it to the specific sites known as Base Station. The given Figure 1 expresses the different components of the WSN system.

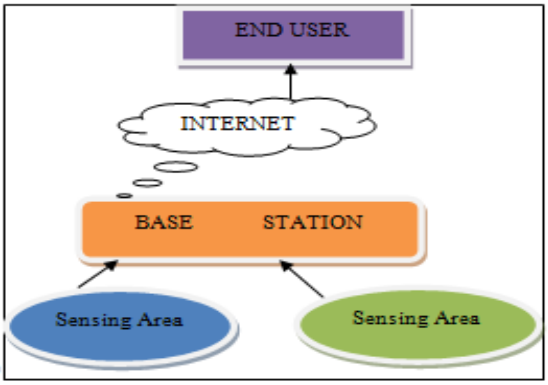

Figure 1: Components of WSN system.

WSN is an emerging field which combines the computation and communication and sensing into a single device. WSN is also having an ability to adapt dynamic changes of the environment. The major research area of the WSN includes the Data Fusion, Routing algorithm, distributed signal processing and security in WSN.

2. EVALUATION METRICES FOR WSN SYSTEM

The WSN systems can be evaluated on the basis of different perimeters. These perimeters include:

- Network lifetime
- Cost
- Temporal accuracy
- Ease of deployment
- Response time
- Security

A. **Network lifetime:** Each node in the network must be designed to manage the local supply of energy to maximize the lifetime of the network.
B. **Coverage:** The coverage of the WSN nodes is beneficial when the nodes are distributed in all the area to be monitored.
C. **Scalability**: There should be no change in the performance whenever in future the new nodes are added to the network.
D. **Response time**: Response time should be fast whenever the WSN setup is for alarm oriented application

scenarios such as Fire Detection.

E. **Temporal accuracy**: The WSN nodes must be able to cross co relate the samples in time, so that exact measure should be forwarded to the end user.
F. **Security**: One of the important measure of the WSN network is security when the WSN is installed for security oriented applications, the security becomes important.
G. **Cost:** The overall cost of the deployment is always demanded in order to maximize the use of WSN for the monitoring of daily events.

3. INDIVIDUAL NODE EVALUATION METRICES

A node can be evaluated on its individual perimeters. On the basis of these characteristics, a node can be assigned different roles e.g. Leaf node, Cluster head node, Associate node etc.Basic components[3] of the node are given in Figure 2 .

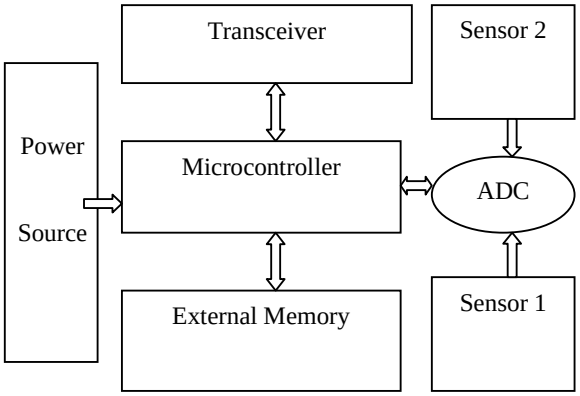

Figure 2: Basic components of a sensor node

The perimeters of a node can be given as follows:

- Flexibility
- Robustness
- Security
- Communication
- Computation
- Size and cost

a) **Flexibility:** The node architecture must be flexible and adaptive. A node must be able to adapt wide range of applications.
b) **Robustness:** In order to support lifetime requirements, each node must be as much as robust. As a node is always demanded to work for years, it should be able to tolerate individual failures.
c) **Security:** A node at individual level must be able to handle and perform encryption and authenticating operations.
d) **Communication:** Key evaluation metric of the node is its communication. The communication range of the node should be high which can ultimately enhance the performance.
e) **Computation:** The CPU of the node must be able to fast decode and compute they arrived data to increase the communication between the nodes at fast rate.
f) **Size and Cost:** The physical size and cost have a great impact on ease and cost of deployment. Physical size of the node also impacts the overall deployment of the node.

4. SENSOR NETWORK CLUSTERING

WSN base station always needs to generate an aggregated value to the end users and the aggregation of the data to be forwarded can also help in reducing the transmission overhead and the energy consumption. To support the data aggregation in the network the nodes can be accommodated in the small groups called the Clusters. Clustering can be defined as the division of the nodes in the groups on the basis of some mechanism. Clustering has been shown to improve network lifetime, a primary metric for evaluating the performance of a sensor network [2]. Clustering is done to achieve the energy efficiency and the scalability of the network. Formation of the cluster also involves the assigning the role to the node on the basis of their perimeters. The coordinator of the cluster which is responsible for the processing, aggregation and transmission of the data to the base station is called the Cluster Head (CH) or the leader, whereas the other nodes which are responsible for sensing and forwarding the collected data to the CH are called the Member Nodes. Figure 3 [6] represents the basis hierarchy of Clustering:

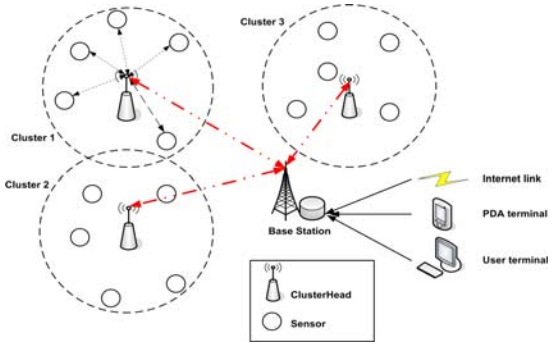

Figure 3: Clustering in WSN

In clustering the 2 tier hierarchy is adopted where in first phase the member nodes sense the data and forward to the CH and in second phase the CH aggregates and process the data to deliver it to the Base Station. The CH node looses more energy as compared to the MN because it performs the fusion on the entire collected data and sends that aggregated report to the BS located far from the cluster location. In a cluster organization both the Intra-cluster and the Inter-cluster communication takes place.
Clustering in WSNs involves grouping nodes into clusters and electing a CH such that:
• The members of a cluster can communicate with their CH directly.
• A CH can forward the aggregated data to the central base Station through other CHs [2].

A. Perimeters of the clustering

- Cluster count/Number of clusters [7]
- Cluster size uniformity
- Inter-clustering routing [7]
- Intra-clustering routing

1. **Cluster count:** On the basis of cluster count the network can be divided into two categories: fixed and variable. Fixed cluster count is that in which the number of clusters in the network are fixed whereas in variable sizes network the number of clusters is not fixed.
2. **Cluster size uniformity:** Cluster size uniformity deals with the size of cluster. It is of two types: Even and Odd. In even cluster size the number of nodes is same in all the clusters of the network and in odd uniformity the cluster size is different.
3. **Inter-cluster Routing:** Inter-cluster routing describes the communication mode of the different cluster. It can be of two types: Single hop and multi hop. Single hop is that type in which the CH communicates with the BS directly. In multi hop clustering the CH communicates with the BS through various intermediate CHs.
4. **Intra-cluster Routing:** It describes the mode of communication between the member nodes and the CH.It can be of two types: single hop and multi hop. In single hop the MN directly with the CH whereas in the multi hop the MN don't directly deal with the CH.

B. Attributes of CH

a) CH coordinates the activities of the cluster.
b) CH is responsible for final transmission of the monitored data.
c) There must be at least one CH in the cluster.
d) The CH can be mobile or stationary .A stationary CH is fixed and easy to maintain whereas the mobile CH is not.
e) The CH can also work as a BS and can take the actions for the cluster.

C. Advantages of Clustering

- Scalability
- Data aggregation
- Less load
- Reduced energy consumption
- Collision Avoidance
- Load Balancing
- Fault tolerance
- QoS

a) ***Scalability:*** As the node is divided into various assignment levels, it makes it easy to add new nodes to the cluster.

b) ***Data aggregation***: Data aggregation helps in reducing the redundant data collected from member nodes.

c) ***Fewer loads:*** Aggregated data avoids the load of the transmission of data from CH to the BS.

d) ***Less energy***: the energy is used less when only non redundant and aggregated data is to be transferred.

e) ***Collision*** Avoidance: By dividing the resources orthogonally to each cluster can leads to a collision free data transmission.

f) ***Load Balancing:*** Equal sized cluster adapt the prolonging of network by balancing the load and prevents from premature energy exhaustion.

g) ***Fault tolerance***: Whenever a node suffers from energy depletion the reclustering can be done.

h) ***QoS***: Clustering protocol helps in delivering a quality and non redundant data to the end user.

D) Design challenges of clustering:

WSN networks become more challenging when they are implemented .The design goals of the WSN are targeted more as compared to wired networks. The WSN is divided into groups called clusters in order to prolong the life of the network .Some of the design considerations in designing Clustering algorithm are[5]:

- Storage
- Security
- Communication
- Limited energy
- Network Lifetime
- QoS

a) ***Storage***: The storage in sensors is very limited and hence it is required to satisfy the storage constraints and query requirement.

b) ***Security***: WSN is very vulnerable to the threats and security. Hence while designing clustering protocols security measures must be added to it.

c) ***Communication:*** communication over the entire region can increase the reliability and also ensures the coverage of the network, to generate the true reports.

d) ***Limited Energy***: Sensor networks are limited by the energy .Clustering can reduce the energy consumptions as compared to the direct communication.

e) ***Network Lifetime***: The limited energy can lead to the reduction in lifetime. Clustering can reduce the energy consumptions by implementing Intracluster communication and multi hop routing schemes.

f) ***QoS***: Clustering always focus on energy efficiency but does not pay attention towards the quality. Hence it is always required to generate a quality in clustering algorithm.

D) Clustering classification:

The clustering algorithms can be divided on various parameters. On the basis of CH rotation for a round it can be synchronous and asynchronous[7].Distributed clustering algorithms can be divided into grid based schemes, PSO based, hierarchical schemes, heuristic scheme .On the basis of data movements it also can be divided into single and multi hop clustering algorithms. The classification of the clustering is given in the figure.[4]

Table 1: Classification of Clustering Algorithms

Sr.No	***Category***	***Algorithms***

1	*Grid Based*	*PEGASIS,PDCH,GROUP*
2	*Hierarchical Scheme*	*LEACH,HEED,EECS,EEUC*
3	*Heuristic schemes*	*LCA-1,LCA-2,HIGHEST CONNECTIVITY*

Grid based : In a grid-based topology, the network is divided into various grids by geography approach. Thus, Grid-based routing generally belongs to location-aware routing [8].The major feature of the Grid based routing is that, the routing operation is performed without maintaining any routing table. Once the BS position is known, the routing operation is performed locally.

Common examples of Grid Based clustering involve:

1) PEGASIS:Power efficient Gathering In Sensor Information System
2) PDCH: Pegasis Algorithm Improving Based on Double Cluster Head

Hierarchical Schemes: The main target of hierarchical routing or is to maintain the energy level of the clusters by using multi hop paths. Hierarchical schemes play an important role in conserving the energy used in the transmission of data. Energy consumed by the sensor nodes for communicating data from sensor nodes to the base station is the crucial cause of energy depletion in sensor nodes [4].Hierarchical Schemes follows the concept of cluster head n leaf nodes. Major algorithms which follow this scheme are:

1) HEED: Hybrid Energy-Efficient Distributed clustering
2) LEACH:Low energy Adaptive Clustering Hierarchy
3) EECS:Energy Efficient Clustering Scheme
4) EEUC: Energy Efficient Unequal Clustering

Heuristic schemes: A heuristic algorithm is an algorithm that usually has one or both of the following goals in solving a problem [9]:

• Finding an algorithm with reasonable run-time (time Needed to set up clusters is affordable); and/or
• With finding the optimal solution

A heuristic search leads to the better performance and not depends on particular perimeters. Some of the algorithms which follow the heuristic search are:

1) LCA-1: Linked Cluster Algorithms version 1
2) LCA-2: Linked Cluster Algorithms version 2

VI. CONCLUSION

Node clustering is very useful in the communication for reducing the overhead of the transmission. In this paper we have only focused on various parameters of WSN and node clustering. There are many challenges on which researchers needs to focus like CH rotation and replacement, Intercluster and Intracluster communication to enhance the lifetime of the network

VII. REFERENCES:

[1] S. Manju Priya, A. Jeevarathinam, "A Study on Energy Routing Algorithms Using Clustering in Wireless Sensor Networks",IJARCSSE, Volume 5, Issue 6, June 2015
[2] Ossama Younis, Marwan Krunz, and Srinivasan Ramasubramanian," Node Clustering in Wireless Sensor Networks: Recent Developments and Deployment Challenges",IEEE,June 2006
[3] Swati Sharma Dr. Pradeep Mittal," Wireless Sensor Networks: Architecture, Protocols", IJARCSSE, Volume 3, Issue 1, January 2013
[4] Vinay Kumar, Sanjeev Jain and Sudarshan Tiwari," Energy Efficient Clustering Algorithms in Wireless Sensor Networks: A Survey", IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 5, No 2, September 2011
[5] Priya Vyas ," Survey on Clustering Techniques in Wireless Sensor Network" (IJCSIT) I, Vol. 5 (5) , 2014,
[6] D. J. Dechene, A. El Jardali, M. Luccini, and A. Sauer.," A Survey of Clustering Algorithms for Wireless Sensor Networks",
[7] Harpinder Kaur, Navjot Kaur," Comparative Analysis of Clustering Protocols for Wireless Sensor Networks", *International Journal of Computer Applications ,Volume 115 – No. 1, April 2015*
[8] Xuxun Liu," A typical Hierarchical Routing Protocols for Wireless Sensor Networks: A Review", IEEE SENSORS JOURNAL, VOL. 15, NO. 10, OCTOBER 2015
[9] D. J. Dechene, A. El Jardali, M. Luccini, and A. Sauer, A Survey of Clustering Algorithms for WirelessSensor Networks"

