DESIGN OF NODE CLUSTERING IN WIRELESS SENSOR NETWORKS

Rajeeb Sankar Bal*

Department of Computer Science and Engineering
ABIT, Cuttack, Orissa – 753014, India
E-mail: rajiv.s.bal@gmail.com
*Corresponding author

Amiya Kumar Rath
Department of Computer Science and Engineering
Dhaneswar Rath Institute of Engineering and Management Studies,
Tangi, Cuttack, Orissa – 754022, India
E-mail: amiyaamiya@rediffmail.com

Abstract: In the wireless sensor networks, hundreds and thousands of wireless sensors are dispersed that collects and transmit data. Also in these networks cluster heads are elected out of the sensors to transmit the data collected to base station. We propose a clustering approach which organizes the whole network into a connected hierarchy and discuss the design rationale of the different clustering approaches and design principles. Further, we propose several key issues that affect the practical deployment of clustering techniques in wireless sensor network applications.

Keywords: WSN, SN, Base station, Clustering.

References to this paper should be as follow : Bal, Rajeeb Sankar., Rath, Amiya Kumar (2013**)** “Design of Node Clustering in Wireless Sensor Networks”, *Int. J. Intelligent Computing and Applied Sciences*,Vol.1, Issue. 1, pp 3-10

1 Introduction

A WSN is a collection of sensor nodes (SNs) organized into a cooperative network. It is clear that ‘smart’ environments represent the next evolutionary development step for building, utilities, industrial, home, transportation, and agriculture. Thus, the interest in WSN is steadily growing. Also WSN consists of a number of sensors spread across a geographical region. In generally, a WSN consists of a host or “gateway” that communicates with a number of wireless sensors (or sensor nodes) via a radio link. Data is collected at the sensor node, compressed, and transmitted to the gateway directly. If the data is required then uses the sensor node (SN). Otherwise, sensor nodes to forward data to the gateway. The gateway then ensures that the data is input into the system. Each wireless sensor is considered a node which presents wireless communication capability, along with a certain intelligence for signal processing and networking data. Depending on the type of application, each node can have a specific address. That is shown in Fig.1.

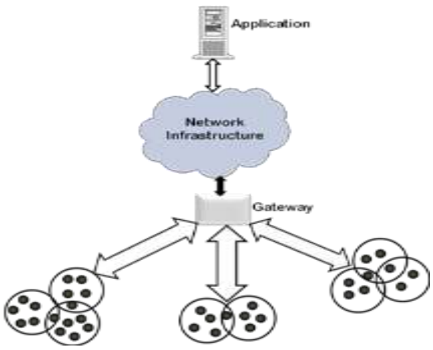

Fig. 1: Sensor in WSN.

WSN targets four main goals:

a. Reading the value of some parameters at a given location and transmitting it to the main processing center. In agricultural environments such as the herd of cattle mentioned earlier, reading the temperature of each cow helps to determine which cow needs closer monitoring.
b. Monitoring the occurrence of certain events such as in a medical application where the peak of the blood pressure and pulse, along with the heart rate, are monitored.
c. Tracking movement of specific objects is widely used in the military to track enemy vehicles.
d. Help classify detected objects, especially in the traffic control environment.

Basic features of sensor networks are:

a. Self-organizing capabilities
b. Short range broadcast communication and multi-hop routing.
c. Dense deployment and cooperative effort of sensor nodes.
d. Frequently changing topology due to fading and node failures.
e. Limitations in energy transmit power, memory, and computing power.

A simple classification of Wireless sensor networks based on their mode of functioning

a. Proactive Networks
b. Reactive Networks
c. Hybrid Networks

A sensor node was developed in North America. A sensor node also known as a “mote”. It is a node in a wireless sensor network that is capable of performing some processing, gathering sensory information and communicating with other connected nodes in the network. A mote is a node but a node is not always a mote. There are two types of sensor nodes used in the WSN. One is the normal sensor node deployed to sense the phenomena and the other is gateway node that interfaces sensor network to the external world. A sensor node consists of five main parts. The one or more sensor nodes gather data from the environment. The central unit, a microprocessor, manages the tasks. A transceiver communicates with the environment and a memory is used to store temporary data or data generated during processing. Data processing tasks are often spread over the network, i.e. nodes co-operate in transmitting data to the sinks (Verdone et al., 2008). The battery supplies all parts with energy, and energy efficiency is crucial. Although most sensor nodes have a traditional battery, there is some early stage research on the production of sensors without batteries, using technologies similar to passive RFID chips shown in figure in 2.

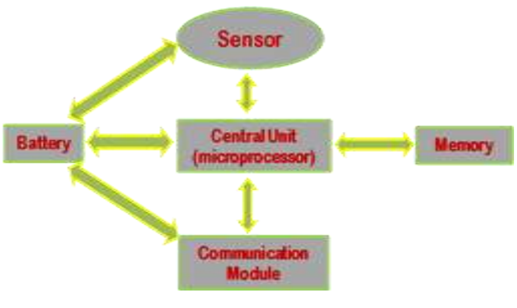

Fig. 2: The components of sensor nodes.

2. Designing of Nodes

The sensor node was to have an x and a y coordinate the idea of using a structure. The structure members consisted of the node identification number as well as the x-coordinate and y-coordinate shown in figure.3.

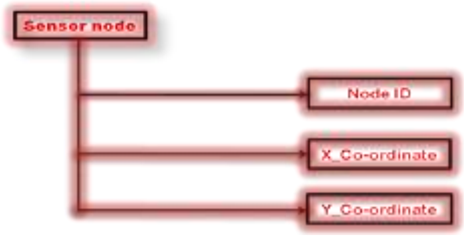

Fig. 3: The Design or Structure of Node.

To get an ID number and set of coordinates for each of the one hundred sensor nodes, the use of a *for* loop was used. Also each node and its information were stored in an array of the structure.

```
for (i=1 ; i<=no ; i++)
{
x_coordinate = Math.random() * x_co;
y_coordinate = Math.random() * y_co;
nodeID[i] = i;
x[i] = xco;
y[i] = yco;
}
```

In the *for* loop “no” used as number of nodes in the network area. Also, the coordinates are coded to exceed the network area which is 10m x 10m. The results of this coding are displayed with two figures in simulation result. The first is the output of the program showing the node ID as well as the random generated x and y coordinates.

3. Design of Node Clustering

In 1978, the technology components for a DSN were identified in a Distributed Sensor Networks workshop on sensors (acoustic), communication, processing techniques and algorithms (including self-location algorithms for sensors) and distributed software. Distributed acoustic tracking was chosen as the target problem for demonstration.

In 1980, the sensor networks was started with the Distributed Sensor Networks (DSN) program at Defense Advanced Research Projects Agency (DARPA) where ARPANET (predecessor of the Internet) approach for communication was extended to sensor networks. The network was assumed to have many spatially distributed low cost sensing nodes that collaborate with each other but operate autonomously, with information being routed to whichever node can best use the information.

Clustering a set of points into a few groups is frequently used for statistical analysis and classification in numerous applications, including information retrieval facility location, data mining, spatial data bases,

data compression, image processing, astrophysics, and scientific computing[1]. Clustering allows data aggregation and limits data transmissions. In the efficient network organization, nodes can be partitioned into a number of small groups called clusters. Each cluster has a coordinator and referred to as a cluster head, and a number of member nodes. Clustering results in a two tier hierarchy in which cluster heads (CHs) form the first tier while member nodes form the second tier (or final tier) shown in figure 4 describes data flow in a clustered network. The member nodes report their data to the respective CHs. The CHs aggregate the data and send them to the central base through other CHs. Because CHs often transmit data over longer distances, they lose more energy compared to member nodes. The network may be re-clustered periodically in order to select energy-abundant nodes to serve as CHs, thus distributing the load uniformly on all the nodes. Besides achieving energy efficiency, clustering reduces channel contention and packet collisions, resulting in better network throughput under high load.

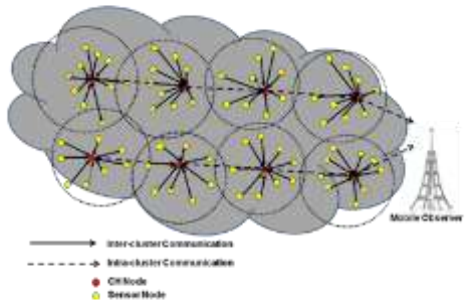

Fig.4: The Data Flow in a clustered network.

Clustering has been pointed out to improve network lifetime, a primary metric for evaluating the performance of a sensor network. Although there is no unified definition of "network lifetime," as this concept depends on the objective of an application, common definitions include the time until the first/last node in the network depletes its energy and the time until a node is disconnected from the base station. In studies where clustering techniques were primarily proposed for energy efficiency purposes (e.g., [2, 3]), the network lifetime was significantly prolonged.

Clustering has been extensively studied in the data processing and wired network literatures. The clustering approaches developed in these areas cannot be applied directly to WSNs due to the unique deployment and operational characteristics of these networks. Specifically, WSNs are deployed in an ad-hoc manner and have a large number of nodes. The nodes are typically unaware of their locations. Hence, distributed clustering protocols that rely only on neighbourhood information are preferred for WSNs (however, most studies in this area still assume that the network topology is known to a centralized controller). Furthermore, nodes in WSNs operate on battery power with limited energy. Hence, the employed clustering approach must have low message overhead. Finally, harsh environmental conditions result in unexpected failures of nodes. Hence, periodic re-clustering is necessary in order to heal disconnected regions and distribute energy consumption across all nodes. Periodic re-clustering is also necessary, as the parameters used for clustering (e.g., the remaining energy, node degree, etc.) are *dynamic*. The clustering techniques proposed for data processing typically consider *static* parameters, such as the distance between the nodes, and assume that nodes are more reliable.

Classification Of Clustering Techniques in Wsns

Clustering in WSNs involves grouping nodes into clusters and electing a CH such that:

a. The members of a cluster can communicate with their CH directly.

b. A CH can forward the aggregated data to the central base station through other CHs.

Thus, the collection of CHs in the network forms a *connected dominating set*.We classify the clustering techniques based on two criteria:

a. The parameter(s) used for electing CHs.

b. The execution nature of a clustering algorithm (probabilistic or iterative).

3.1. Node Clustering

The different categories of clustering algorithms and recent efforts to design clustering methods for various kinds of graphical data. Clustering algorithms are typically of two types. The first type consists of node clustering algorithms in which we attempt to determine dense regions of the graph based on edge behavior. The second type consists of structural clustering algorithms, in which we attempt to cluster the different graphs based on overall structural behavior shown in figure.5.

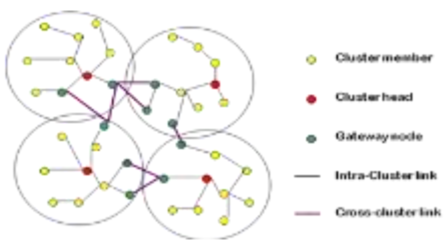

Fig: 5 Node Clustering Structure.

Clustering mechanisms have been applied to sensor networks with hierarchical structures to enhance the network performance while reducing the necessary energy consumption [4]. Clustering is a cross-cutting technique that can be used in nearly all layers of the protocol stack. The primary idea is to group nodes around a cluster head that is responsible for state maintenance and inter-cluster connectivity.

In clusters without any cluster head, a proactive strategy is used for intra-cluster routing while a reactive strategy is used for inter-cluster routing. However, as the network size grows, there will be heavy traffic overhead within the network [5]. Therefore, normally one node is selected as the cluster head of a cluster, and it acts as the local coordinator of transmissions within its cluster. A hierarchical routing or network management protocol can be more efficiently implemented with cluster heads. As compared to the base stations used in current cellular. Systems, the cluster head does not have any special hardware, and is in fact dynamically selected among the set of nodes. However, a cluster head performs additional functions as a central administration point, and a cluster head failure would degrade the performance of the entire network. It may be the bottleneck of the cluster. An efficient node clustering mechanism tends to preserve its structure when a few nodes are moving and the topology is slowly morphing. The objective of the node clustering procedure is to find a feasible interconnected set of clusters that covers the entire node population. For the initial deployment of the network, the nodes could be deployed in the coverage area regularly or randomly.

4. Simulation

4.1. Designing of Node

The result of designing node consists of nodeID as well as the random generated x and y coordinates.

INPUT:

Enter maximum number of node (000 - 200): 10
Enter number of x_coordinate (000 - 200): 10
Enter number of y_coordinate (000 - 200): 10

OUTPUT:

Network Area = 10m x 10m

Node ID:	X-Coordinate:	Y-Coordinate:
1	0.105594	0.039674
2	3.351543	0.332652
3	3.557237	2.172002
4	5.369732	1.957762
5	7.003387	9.499191
6	2.747887	4.442884
7	1.089816	6.982330
8	5.643483	0.415052
9	1.651662	8.154851
10	6.855373	7.643971

Fig: 6 The NodeID, x and y co-ordinate of Sensor Nodes.

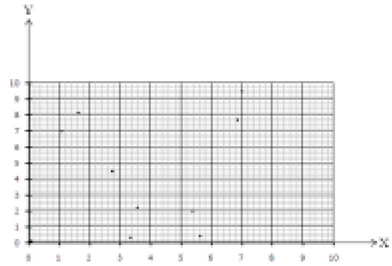

Fig. 7: The Nodes randomly generated inside of the network with range of 10mX 10m.

4.2. Clustering Approaches

A graph G is a pair of sets (V, E), where V is set of vertices and E is the set of edges. If G is a directed graph (digraph) the elements of E are ordered pair of vertices. In this case an edge (u,v) is said to be from u to v and to join u to v. If G is non-directed graph the elements of E are unordered pairs (sets) of vertices. In this case an edge {u,v} is said to join u and v or to be between u and v. In graph theory, a dominating set for a graph G = (V, E) is a subset S of *V* such that every vertex not in S is adjacent to at least one member of S. The domination number $\gamma(G)$ is the number of vertices in a smallest dominating set for G.A subset S of sensor nodes is called a dominating set of G, if every node in V - S is connected to some node in S. The minimum number of nodes in a dominating set of G is called domination number of G as shown in figure 8.

Definition adjacency matrix

A representation of a directed graph with n vertices using an n × n matrix, where the entry at (i,j) is 1 if there is an edge from vertex i to vertex j; otherwise the entry is 0. An undirected graph may be represented using the same entry in both (i,j) and (j,i).

Domination set problem

We describe an algorithm for the domination set problem to find the cluster heads in a cluster group. Encouraging node lies outside the dominating set and is adjacent to maximum number of cluster heads (CH). Encouraging node of G is a node v, where v ∈ V−S, such that v is connected to maximum number of nodes in S.

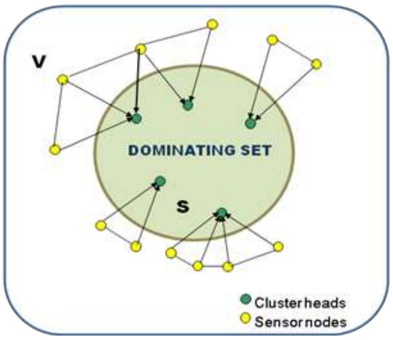

Fig. 8: The Dominating set based cluster head.

ALGORITHM:" *Algorithm for Election of Cluster heads using Dominating Set".*

Step No:

1. Using adjacency matrix to input the sensor network graph G.
2.Create a vector to count the row sums.
3.Consider a vertex v with the highest degree in the graph G and put v into the dominating set S.
4. Eliminate all the edges incident to v in the graph and update the row sums.
5. If G has no edges , STOP. Go to step 5. Else go to step 3 (G has no edges, if all the entries of the adjacency matrix of G are zeros).
6. The remaining vertices of G form a dominating set of G

Cluster formation Using Adjacency algorithm

The base station (BS) dynamically initializes the cluster heads (CH) to from their cluster groups by sending a broadcast packet i.e., message. The cluster heads send a broadcast at its maximum transmission range. All nodes that hears the broadcast generates a node database containing node ID, and number of other CH's sharing the same node. All sensor nodes hearing the broadcast from cluster head will send an acknowledgement message tagged with the node data base. Cluster heads decides the members of the group on receipt of the acknowledgement messages. After cluster groups are formed, the time synchronization signals will be broadcasted from the Encouraging node to the cluster members through its respective cluster heads. We present cluster formation by the adjacent vertices sensor nodes to the highest degree vertex cluster head and cluster groups can be formed as shown in figure 9.

1. To determine cluster heads from dominating set S of the graph G = (V, E).
2. Enter an adjacency matrix.
3.Consider a cluster head and put all its neighbours in its Cluster. The neighbours of a node correspond to '1' in the respective row of the adjacency matrix.
4. If there is no more cluster heads, STOP. Otherwise, go to the next cluster head and go to step.3

.

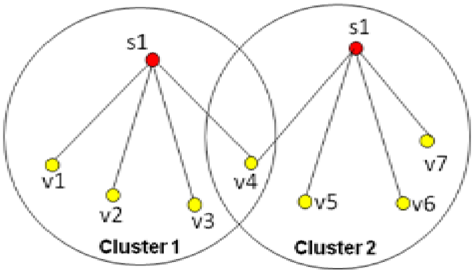

Fig: 9 The Cluster formations from adjacency matrix

In the sensor network, the sensor nodes are denoted by v1, v2, v3, v4, v5, v6, and v7 and cluster heads by s1 and s2.The row sum gives the degree of each vertex and the highest degree vertices would be the cluster heads.

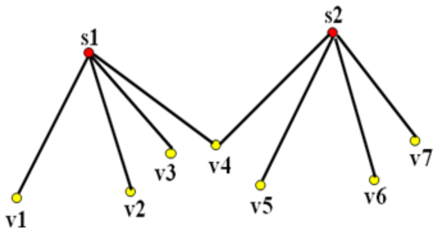

Fig: 10 The Graphical Representation of Sensor Node and Cluster Head

RESULT:
Input Number Of Vertices :3
Input Adjacency Matrix
1
0
0
1
1
0
1
1
1
Inputed Adjacency Matrix :
 1 0 0
 1 1 0
 1 1 1
Degree Of Each Node :
V0---->1
V1---->2
V2---->3
Maximum Of Degree =3
Clustered Head Node = V2

5. Conclusions And Future Work

Here, we introduced, to simulate and design sensor nodes in WSN. The program created was simulation of a network the ranged from Nm x Nm, where N is number of sensor nodes in WSN (N=1 to 1000 or more), but asked the user for the number of nodes to be dispersed into the sensor field. Node clustering is a useful topology management approach to reduce the communication overhead and exploit data aggregation in sensor networks. We have designed the clustering approaches according to the clustering criteria and the entity responsible for carrying out the clustering process. In the future, we would study the impact of transmission range on the connectivity level of the sensors to cluster heads.

Reference

[1] Exact and Approximation Algorithms for Clustering (Extended Abstract). Pankaj K Agarwal, Cecilia M. Procopiuc-1998.

[2] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "An Application-Specific Protocol Architecture for Wireless Microsensor Networks," IEEE Trans. Wireless Commun., vol. 1, no. 4, Oct. 2002, pp. 660–70.

[3] HEED: A Hybrid, Energy-Efficient, Distributed Clustering Approach for Ad-hoc Sensor Networks Ossama Younis and Sonia Fahmy Department of Computer Sciences, Purdue University 250 N. University Street, West Lafayette, IN 47907–2066, USA.

[4] V.Mhatre and C. Rosenberg, "Design guidelines for wireless sensor networks: Communication, clustering and aggregation", Ad Hoc Networks, vol. 2, no. 1, Jan. 2004,pp. 45 – 63.

[5] V.Mhatre and .Rosenberg, "Homogeneous vers heterogeneous clustered sensor networks: A comparative study", in Proceeding of IEEE ICC' 04, vol. 6, Paris, France, June 2004, pp. 3646–3651.

Biographical Notes : **Mr. Rajeeb Sankar Bal is working as** Senior Lecturer in the Department of Computer Science & Engineering of Ajay Binay institute of Technology at Cuttack, Orissa, India. He received his MCA degree from Fakir Mohan University and MTech in Computer Science & Engg from BPUT, Odisha. His research interests include Sensor Network, Soft Computing

Prof (Dr) Amiya Kumar Rath is working as professor in the department of Computer Science and Engineering of Dhaneswar Rath Institute of Engineering & Management Studies, Cuttack, Orissa, India. The author has obtained his B.E. degree in Computer Science & Engg. in the year 1990, MBA degree in Systems Management in the year 1993, M.Tech in Computer Science in the year 2001 and Ph.D in Computer Science in the year 2005 from Utkal University for the work in the field of Embedded system and Served as professor in different premier institutes since 20 years. He published over 50 papers and 4 books from top international publishing house. He has supervised ten PhD theses in different areas of machine intelligence. His current research interest includes Embedded System, Adhoc Network, Sensor Network, Power Minimization, Biclustering, Evolutionary Computation and Data Mining.