



Survey on Cluster Based Data Aggregation in Wireless Sensor Network

Shivani Desai

Department Of Computer Engineering
Institute of Technology, Nirma University
Ahmedabad, India

Anuja R. Nair

Department Of Computer Engineering
Institute of Technology, Nirma University
Ahmedabad, India

Tarjni Vyas

Department Of Computer Engineering
Institute of Technology, Nirma University
Ahmedabad, India

Hunar Jain

Department Of Computer Engineering
Institute of Technology, Nirma University
Ahmedabad, India

Abstract: Different data aggregation protocols are required in wireless sensor network (WSN) to reduce computational and communication power consumption of various nodes to enhance battery lifetime of the nodes. Recent advances in WSN have led to many new protocols specifically designed to conserve the energy of the nodes and increase the lifetime of the nodes. This paper presents an overview of what WSN is and why data aggregation is required and it also focuses on reducing the power consumption of WSN by using clustered based approached protocols namely LEACH by taking certain Clustering Parameters into consideration like number of nodes, overlapping, algorithm complexity.

Keywords: WSN, LEACH, Data Aggregation, power consumption

I. INTRODUCTION

Wireless Sensor Networks (WSN) are widely distributed sensors (nodes) to measure change in physical and environmental conditions by sensing variations in temperature sound pressure etc., to achieve this purpose sensor nodes work together in a predetermined fashion to transfer the data to a main location (generally refer to as sink node). Modern networks are bidirectional making it easy for controlling the sensor activity. These nodes have computational and communication capabilities making them a good choice over conventional cables for operating in varied locations and harsh environments. Modern nodes are compact, consume less power and are generally adaptive to conditions they are put in.

A typical design of WSN is shown in the figure where the sink node is collecting data from the sensor nodes and giving it to the server (internet in this case) which in turn providing the data to the user.

A. Basic Characteristics of the WSN

- Limited power consumption as sensor nodes works on batteries.
- Nodes are mobile opposed to conventional wires.
- Usage in harsh environment.
- Nodes can be deployed in any fashion over an area.
- Easy to use and multi-layer design operation.

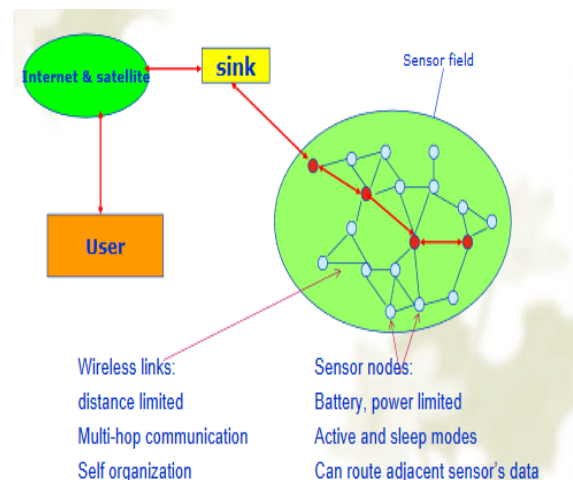


Figure1. Architecture of WSN

B. Advantages of WSN

- No fixed infrastructure is needed nodes can be deployed in any manner.
- Non-Accessible Places like forests, mountains, hilly areas can be suitably reached.
- Cheap and less wiring.
- New devices can be added readily with flexible control due to multi-level protocols.
- Acentralized system can be used to control the data from the nodes working behind the sink node.

C. Disadvantages of WSN

- a) Complicated configuration as compared to wires.
- b) Energy/Power Consumption should be less since they are battery operated they have a limited lifetime.
- c) Range of nodes is critical while deciding connectivity with the sink and other nodes, it limits the operating area and the number of nodes to be used.
- d) Hundreds of thousands of nodes are required so they have to be cheap.
- e) Duplication of data is a major issue, which will waste energy and resources.
- f) Limited storage of node so limited amount of data can be stored and transmitted further.

a single One. Intermediate nodes can be there between primary sensor and header node.

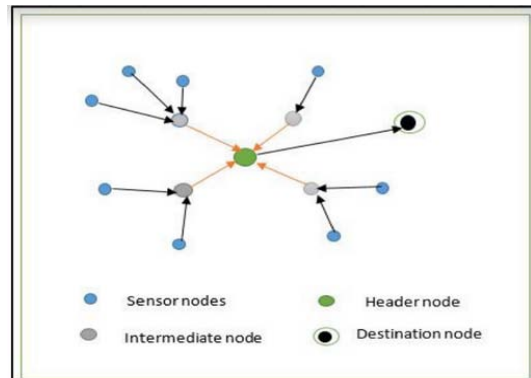


Figure 2. Centralised Data Aggregation [1]

D. Applications of WSN

- a) Area Monitoring: - WSN is deployed over a region to be monitored especially in case of military related purposes.
- b) Earth Sensing: - To calculate pressure and temperature of various regions having harsh conditions.
- c) Air Pollution Monitoring
- d) Health Care Monitoring: - Wearable and implanted nodes are used on the patient or at its close proximity which can further used to measure body position and other stimulation.

b. In-Network Approach: It is of two types :-

1. With Size Reduction: - Data received or combined from various nodes is compressed using various methods and algorithms to reduce the power consumption involved.
2. Without Size Reduction: - It only includes aggregating the data from neighbors and converting into a single packet for further transmission.

II. DATA AGGREGATION

To avoid duplication and save resources (energy and bandwidth) data aggregation is used. In a Wireless Sensor Network number of nodes collect information and this information is transferred to base station where it is processed in doing so high amount of battery consumption take place, at the same more bandwidth is being consumed to avoid this data is processed at various nodes before giving it to the sink node such processing is called Data Aggregation.

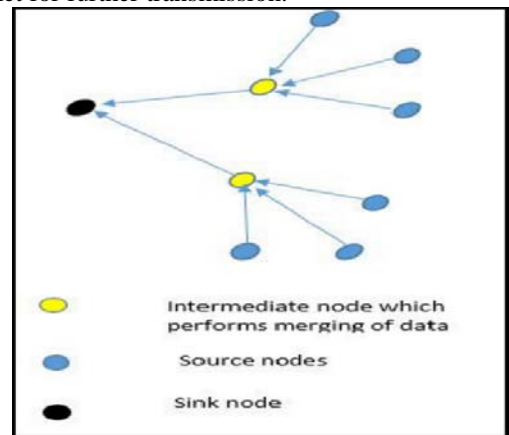


Figure 3. In-Network Data Aggregation [1]

A. Features of Data Aggregation

- a) Increasing network lifetime by using resources in an eventful manner using various algorithms
- b) Using Shortest Path algorithms nearest node is reached.

B. Techniques of Data Aggregation

There are four techniques to achieve Data Aggregation namely:

- a) Centralized Approach
- b) In-Network Approach
- c) Tree Based Approach
- d) Cluster based Approach

a. Centralized Approach:- In this method, each sensory node selects the most powerful node among its neighbors and sends it the data packets via the shortest possible route. This powerful node is commonly known as header node which is capable of performing data aggregation and combining the data various packets into

c. Tree Based Approach: - A tree is formed first known as DAT (data aggregation tree), for each transmission, minimum spanning tree is created. Each node acts as leaf send its data to parent node, which acts as branch node, and data in this way flows from leaf node to sink and the final parent node does the processing.

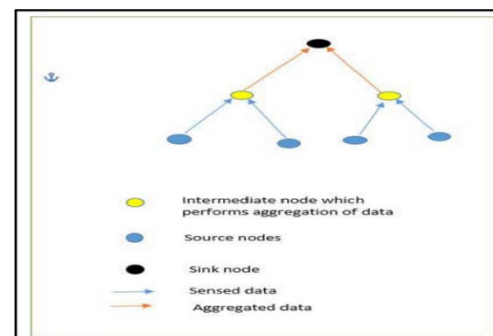


Figure 4. Tree Based Data Aggregation [1]

d. Cluster Based Approach:- Most efficient approach in which network is divided into clusters containing various nodes in each. In each cluster, one node is selected as Cluster Head this done using algorithm and after one process cluster head changes so that optimization can be achieved. These cluster heads perform the aggregation reducing the packet size as well as redundancy and then transmit the packet to the required location.

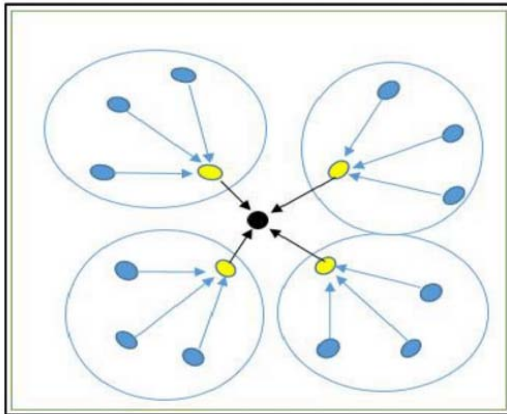


Figure 5. Cluster Based Data Aggregation[1]

C. Comparison between Various Data Aggregation Approaches

Table 1. Comparison between various approaches

Approach	Delay	Redundancy	Accuracy	Traffic	Energy Consumption
Centralised	Moderate	Moderate	Moderate	High	High
In-Network	Moderate	Moderate	Less	High	Moderate
Tree	Less	Less	Moderate	Moderate	Less
Cluster	Less	Less	High	Less	Less

A Comparison between various approaches clearly seems to sum up that cluster approach is a good choice because not only energy consumption is taken into account while writing the algorithms but also the traffic load is reduced in comparison to tree approach as shown in Table 1.

III. CLUSTERING IN WSN

As concluded above clustering technique is far more efficient than the rest. Hierarchical Clustering protocols are extremely useful rather than single-tier communication considering the area where sensors needed to be deployed

and the battery energy of the sensors. Moreover, sensors are deployed in thousands of numbers and the possibility of damage is high so an efficient way of communication is needed in which even if some nodes are compromised communication can be achieved (scalability can be achieved at ease). Grouping Sensor nodes to form cluster is widely accepted in order to fuse data and make aggregation possible leading to saving energy of many intermediate nodes.

A. Advantages

1. Scalable for thousands of nodes.
2. Load balancing and efficient resource utilisation.
3. Reduce the size of routing table at individual nodes.
4. Conserve communication bandwidth and avoid redundant messages among nodes.
5. Nodes can be scheduled to low-sleep power mode by CH to save energy and round robin order can be set to avoid collisions.

B. Working in a cluster

- a. In such networks, each cluster has a leader called as Cluster Heads (CH) who performs aggregation of the data received from several nodes.
- b. Several other sensor nodes in a cluster (SN) are members of the same cluster.
- c. CH nodes form the higher level and SN forms the lower level in the hierarchy of data flow.
- d. CH aggregate the data (decreasing the total relayed packets) and transmit to the base station.
- e. CH nodes spend higher energy compared to SN because data has to be sent over higher distances.
- f. Protocols are used to periodically change the cluster heads in a cluster to save energy.
- g. Base Station(BS) is the final processing point of the information received from various cluster heads.

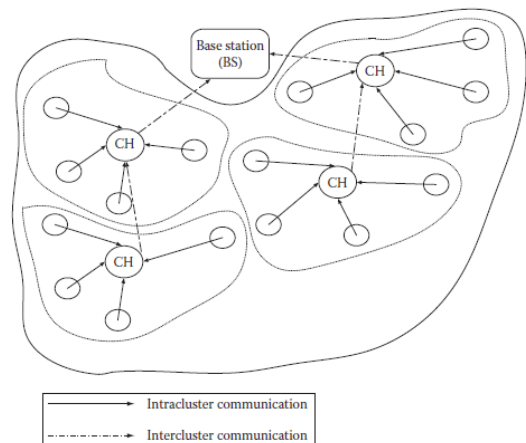


Figure 6. Data communication in clustered network [2]

C. Clustering parameters

In order to design clustering algorithms some designated parameters have to be kept in mind, which will optimize the algorithms to serve best, and these are-

- a. Number of clusters: - Includes CH election and the number of clusters because this number will decide the efficiency of routing protocol.
- b. Nodes and CH mobility: - Includes dynamically changing the membership of cluster nodes.
- c. Node types and roles: - Special computation capabilities to some nodes or not.
- d. Cluster Head Selection:-By probabilistic approach.
- e. Algorithm Complexity:-Fast execution of the protocol is primary goal.
- f. Overlapping: - Minimize the overlapping of nodes of different cluster together.

Based on the above parameters a priori probability is assigned to each sensor node to determine initial CH. The probabilities initially assigned to each node serve as initial criteria in order for the nodes to decide individually on their election as CH. This leads to achieve faster execution time or convergence time and reduced volume of exchanged packets.

IV. Clustering Protocol LOW ENERGY ADAPTIVE CLUSTERING HIERARCHY (LEACH)

Features

- a. First dynamic clustering protocol addressed the WSN needs, using homogenous nodes, which randomly deployed.
- b. Hierarchical, probabilistic distributed one-hop protocol focusing on
- c. To reduce energy consumption by data aggregation at CH's.
- d. To improve lifetime by distributing the resources.
- e. CH nodes as routers to the BS (base station).
- f. Nodes make autonomous decision without centralized control.

A. Working

Initially a node decides to be CH with probability “p” and broadcasts its decision. After its election, each CH node sends a message to sensor nodes so that they can choose the CH, which can be reached with least communication energy (based on the signal strength of the CH's message).

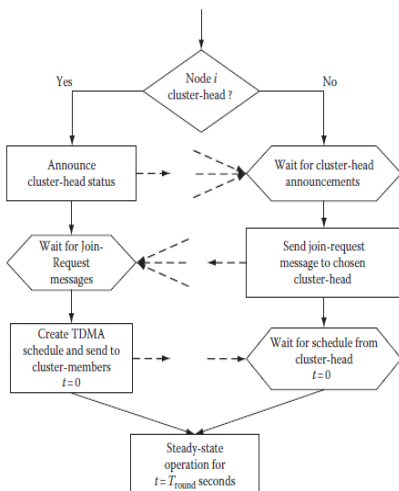


Figure 7. Flowchart of the cluster formation process of LEACH [3]

CH is changed periodically to balance the load. The rotation is performed by getting each node to choose a random number “T” between zero and one. A node becomes a CH if the number is less than the following threshold:-

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

where
p is the desired percentage of CH nodes in the sensor population
r is the current round number
G is the set of nodes that have not been CHs in the last 1/*p* rounds

Figure 8. Equation to choose random number ‘T’. [4]

B. Advantages

- 1. A good balancing of energy by random rotation of CHs.
- 2. Scalability for cluster formation.

C. Disadvantages

- 1. Probability of selection of CH with very low energy.
- 2. It forms one-hop intracluster and intercluster topology where each node should transmit to CH therefore it is not that effective where network area is quite large.

V. CONCLUSIONS

This paper discusses about the introduction to WSN and their present application in various fields. With increasing use of these resources, data aggregation is required, this paper discusses various approaches to it while covering the two most important LEACH protocol based clustering approach and their advantages and disadvantages. As for future work, we will test L.E.A.C.H algorithm in a software and analyze the data from it.

VI. REFERENCES

- [1].Sumedha Sirsikar,Samarth Anavatti,“Issues of Data Aggregation Methods in Wireless Sensor Network: A Survey”,Procedia Computer Science 49 (2015) I94-201.
- [2].Basilis Mamalis, Damianos Gavalas, Charalampos Konstantopoulos,Grammati Pantziou,“Clustering in Wireless Sensor Networks”,Zhang/RFID and Sensor Networks AU7777_C012 Page Proof Page 325 2009-6-24.
- [3].Basilis Mamalis, Damianos Gavalas, Charalampos Konstantopoulos,Grammati Pantziou,“Clustering in Wireless Sensor Networks”,Heinzelman, W.B. et al., IEEE Trans. Wireless Commun., 1, 660, 2002.
- [4].Basilis Mamalis, Damianos Gavalas, Charalampos Konstantopoulos,Grammati Pantziou,“Clustering in Wireless Sensor Networks”,Zhang/RFID and Sensor Networks AU7777_C012 Page Proof Page 335 2009-6-24.