IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 18, Issue 1, Ver. IV(Jan – Feb. 2016), PP 14-18
www.iosrjournals.org

Energy Efficient Query Optimization in WSN usingThreeLevel Modelling

C.Komalavalli[1],VijayBhushanAggarwal[2]
[1]*Department of Information Technology, Jagan Institute of Management Studies, Rohini, New Delhi, India*
[2]*Department of Information Technology, Jagan Institute of Management Studies, Rohini, New Delhi, India*

Abstract:*With advancement in microprocessor technology, wireless sensor networks is playing vital role in different types of applications such as health care monitoring, fire detection etc. In these applications, large amount of data generated by the sensors are communicated to the base station, which will provide final result for the query. Since the sensor nodes are energy constraint devices, energy efficient query methods and storage of the data is required. Clustering is playing an important role in enhancement of the energy level of the sensor and scalability of the network. Since communication depletes the energy level of the sensor, further selection of cluster head is important for reducing the communication cost. In this paper, we are introducing the model along with algorithm for the cluster head selection and efficient query processing in WSN.*
Keywords:*WSN, Clustering, Base Station, Cluster Head, Tiny DB*

I. Introduction

Wireless sensor network is a network which consists of large number of sensor nodes distributed in a geographical area for monitoring temperature, humidity, sound, pressure etc.[1]. The distance between nodes is small compared to wired network. They are able to sense the phenomenon for which they deployed process and send the data to the centralized server known as a gateway. Sensor nodes are battery powered and energy getsdepleted in the nodes in due course of time. It is impractical to charge or replace the battery in the sensor nodes of the WSN. Since the communication from nodes is more energy consuming than the computation at the nodes, it is very important to minimize the communication at the nodes.

Users are sending the queries to the base station which are then transmitted to the sensor nodes. Every time, the sensors are collectingrequired data which then communicated to the base station. For managing these data, WSN is viewed as distributed database and sensors are considered as datasources. They are storing the datain the rows of a relation, same like traditional database systems.WSN can be programmed for data sensing and collection. Since the energy consumption is one of the major issues, data collection process from the nodes to the base station should be optimized so as to minimize the energy consumption. [2]

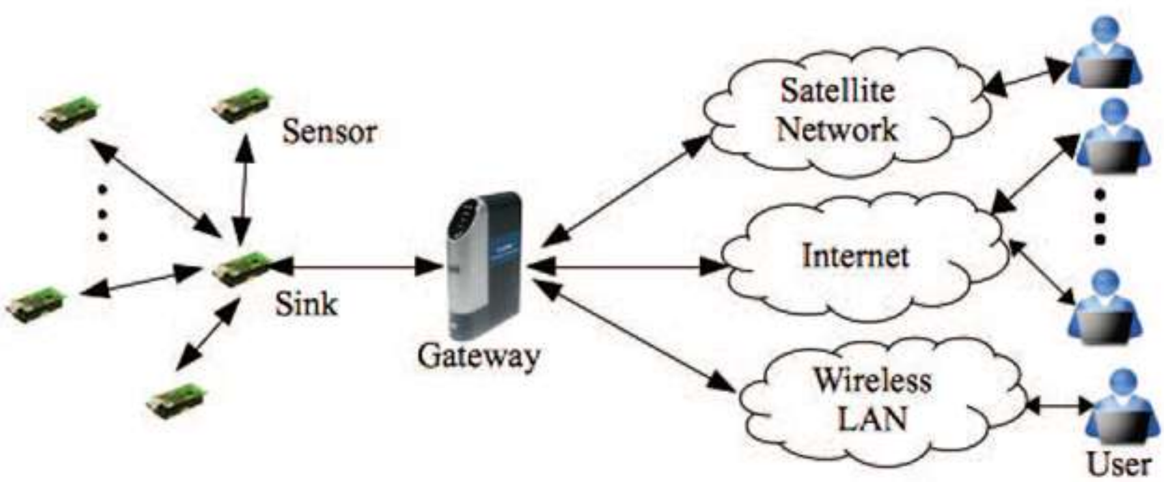

Fig:1 Wireless Sensor Network Architecture[1]

The limited computing power, life and energy of sensor nodes are to be considered for the efficient query processing.

Since energy depletion is mainly due to communication between nodes, various reduction techniques are existing for resolving this issue. Data aggregation techniques aggregates the data in the intermediate nodes and packet merging combines the small pockets into big one without considering the semantics and correlation between packets. Data compression techniques involve data encoding and decoding at the sensor and sink node respectively.

Sensors can be placed anywhere, data should be collected and sent to the base station. In order to minimize the communication cost of query processing, various optimization techniques for WSN at different levels have been discussed in the literature.[1]

Our consideration in this paper focuses on network level and it involves query dissemination, data aggregation, network topology and programming. Queries can be compared according to the query attributes or query conditions.

To reduce the overall cost of query processing, we are considering threelevels in the network: Base station, cluster head and sensor node. Clustering means group of nodes workingtogether under one leader known as cluster head. Cluster heads [3] also one of the sensor nodes and can be selected using different algorithms. Sensors will be sending data to the cluster head which in turn will communicate to the base station, thus reducing the overall communication cost. Cluster Head also maintains the metadata about the sensors in its own cluster and physical location within the network.[4]. The key attributes such as number of clusters, nodes in each cluster, intra cluster communication, nodes role, cluster head selection and multiple level cluster approach are to be considered for the designing the cluster architecture. [5] [6]

The rest of the paper is organized as follows. Section 2 includes related work. Section 3 describes problem formulation. Section 4 describes the proposed architecture and the algorithm optimizing communication cost. Finally Section 5gives conclusions and future scope of this work. Section 6 gives references.

II. Related Work

The sensor databases are different fromour traditional databases. The latter are mainly focussing on collected data rather than extracting the data from the network. But sensor networks are focussing of both the extraction and collection of the data from the network.

The two approaches, data ware house and distributed approach are existing for the efficient data storage and query processing in WSN. In data warehouse approach, the sensors are collecting the data and sending the data to the base station periodically. Only base station is responsible of processing the entire data. This approach leads to the wastage of resources and creating bottleneck because of large volumes of data from sensors to base station.[2]

In the second approach, the node itself is treated as adata source and they are not sending the data periodically to the server, thus it forms a distributed approach.[4]. The queries are injected into the base station and disseminated into network according to the routing protocol. The sensor nodes send their data to their parent node called cluster head and cluster head aggregates the data from different nodes and sends that to the base station. This approach is called network processing and it reduces the amount of transmitted data and latency.[3]

There are so many clustering protocols which are proposed in the literature such as LEACH, HEED, TEEN, PEGASIS, and EEHC for efficient processing of the queries.

Low Energy Adaptive Clustering Hierarchy (LEACH) algorithmwas proposed by Heinzelman et.al [7] in which the cluster formation is in a distributed manner. In this algorithm, any sensor node can be the cluster head randomly and it broadcasts its status to other nodes. Thus a node with low energy can be selected and if that node dies, that particular cluster will become non-functional leading to inaccurate result. This protocol is not focussing on data recovery techniques at the time of node failure.

Power-Efficient Gathering in Sensor Information Systems (PEGASIS) was proposed by S. Lindsey et al. [8] in which each node receives and transmits only to its close neighbours. Every node becomes a leader and transmitting the data to the base station. This approach distributes the load evenly among the sensor nodes.

Manjeshwar et.al [9] introduced an algorithm called Threshold sensitive Energy Efficient sensor Network protocol (TEEN) which is designed for distributed reactive network applications. It is based upon hard and soft threshold values. If these values are not received by the sensor nodes, then the data cannot be retrieved by the network.

Hybrid Energy Efficient Distributed Clustering (HEED) was proposed by Younis et al. [10] for real time applications. It is based upon residual energy and intra-cluster communication cost. Unlike LEACH, only sensors having higher energy can be selected as cluster head. For finding the communication cost, both the intra and inter cluster communication cost is used but it is not considering the data recovery of the network.

Energy Efficient Hierarchical clustering (EEHC) [6] is a randomized, distributed clustering algorithm and that has been divided into two stages. In the first stage, each node itself announces as cluster head and becomes the volunteer on the basis of the probability of the node. In the next stage, clustering hierarchy is formed and aggregated data will be sent using the hierarchy level. The node which is nearest to the base station becomes the relay for the other cluster heads.

III. Problem Formulation

Clustering is the most convenient way of managing the communication overhead and data aggregation in wireless sensor networks. Even though large numbers of algorithms are present but energy consumption during cluster formation is very high. So working condition of the cluster head is playing an important role when the nodes are dying. In this paper, we propose an algorithm for the checking the working condition of cluster head and selection of cluster head according to energy level.

IV. Proposed Architecture

Assumptions:

- Sensor nodes are static in nature.
- Sensors broadcast their energy level to the bases station periodically.

4.1 System Model

The network consists of many clusters and each cluster consists of several sensor nodes. The clusters are denoted by c1, c2, and soon. Every sensor has a unique id and will be denoted by cluster number and sensor number. If the nodes are in first cluster, then id will be c101, c102, and soon. In every cluster, one node will be considered as cluster head and sensor id will be assigned as cluster head id. We are maintaining these details in the base station. Only cluster head can communicate to the base station and send the aggregated data to the base station.

In the base station we are managing two tables namely cluster head details and sensors details.

Table I: Schema for the Cluster Head details

Cluster Head ID	Stores the node ID which is currently Cluster Head
Position	Latitude and Longitude(Lat 40°26'North; Long3°42'West)
Number of Nodes	Number of nodes in that particular Cluster
Status	Working condition of Cluster head (Working or Not working)

Table 2: Schema for the Sensors Details

Cluster Head ID	Stores the node ID which is currently Cluster Head
Sensor ID	It stores the ID of the sensors
Criteria	Phenomenon such as Temperature, Sound, Pressure etc
Data Unit	Celsius,Decibels, Pascal etc.
Energy Level	Energy Level of the Sensor in Joules

4.2 Query Execution

Users will forward the query to the base station. Every query will have unique id.For efficient query processing, Meta data is playing a vital role. Meta data is managed in two levels: 1) Base station and2) Cluster Head.

There are two phases for executing the query. First phase is called setup phase and second phase is called execution phase.

4.3 Setup Phase

In the base station, we are managing the data about cluster head details and details about the sensor nodes also. All sensor nodes broadcast their energy level to the base station and keep updating the sensor table periodically.

Initially, the base station checks whether the result of the query is available in the memorycache. If the result is present, it retrieves and returns the result to the user. If the result is not there, base station checks for cluster head working condition. If all cluster heads are working, then query will be disseminated to all the cluster heads of the network.

If the status of any cluster head is "not working", then base station checks the energy levels of all sensor nodes in that particular cluster. The node which is having maximum energy will be selected as cluster head and this will be communicated to all the sensors in that particular cluster. The id of this sensor will be assigned as cluster head id and updated the cluster head table.

Setup cost in the base station can be calculated as:

$$SC(B_s) = \sum_{i=1}^{n} (h * v) \quad (1)$$

Where n represents the number of clusters, h is number of hops between cluster head and base station and v represents the size of the data. The cost has to be calculated only once when the cluster head is not working in the particular cluster. Then it remains same for further query processing.

4.4Query Execution Phase

In the cluster head, sensor table(Table2) is also maintained and it gets updated alongwith the base station.

In this phase, query will be forwarded from the cluster head to the child nodes using modified Bellman Ford algorithm.[4]

This algorithm was proposed by the authors [4] for getting the shortest path. The algorithmhas been modified here according to the working condition of the cluster head. Cluster head checks for the sensor table and find the sensor type matching with the query. Then query will be disseminated to that particular nodes and data will be retrieved from the nodes. Cluster head collects, aggregates the data and send the result to the base station.

Modified Bellman Algorithm [4]

```
Step 1:
Initialization
For each vertex v in vertices
If v is source then
            v.distance=0;
      Else
            v.distance=infinity;
            v.predecessor=null;
Step 2:
            For i to size [vertices]-1
                  For each edge UV in edges
      If node flag=1 and v.flag=1Then// working condition of cluster head is checking here
            U=uv.source
            V=uv.destination
      If u.distance + uv.weight<v.distance
                  v.distance=u.distance=uv.weight
                  v.predecessor=u
                  Next edge
            Next i
      // for negative cycle
       For each UV in edges
      U=uv.source
      V=uv.destination
      If u.distance + uv.weight < v.distance
                  Then error "graph contains negative weight cycle"
            Loc=0
      V=sink
      While v! =source
      Route [loc++] =v
      If (v.predecessor == source)
                        then query processing will be done
      Delete_edge (v, v.predecessor)
      V=v.predecessor
```

Since within a cluster C_i, nodes are represented as C_iS_1, $C_iS_{2,\ldots\ldots}C_iS_j$,….,.$C_iS_{mi}$,thereforeQuery Execution Cost ($QxE(C_i)$) for cluster C_i, can be calculated as:

$QxE(C_i) = \sum_{j=1}^{mi} \sum_{j=1}^{mi}$ CiSjwhere S_j= projection (d) * size (d) (2)

whered is phenomenon such as temperature, sound, pressure,etc.(matching criteria).Thus, total Query Execution Cost for all the clusters C_1, C_2, …. C_n can be calculated as:

Total QxE=$\sum_{i=1}^{n}$ QxE(Ci) (3)

Once the base station receives the aggregated data, it can process and send the results back to the user. At the same time, data is stored in the base station for future use also.

This model reduces the setup cost by checking the working condition of the cluster head which is maintained in the base station. The status and energy level of all sensorsare maintained in the base station. In case all cluster heads are in working condition, the processing proceeds to the next phase. If any cluster head is not working, then another sensor in that cluster having maximum energy level is selected as new cluster head. This maximizes the life of the newly selected cluster head.

In each cluster, large numbers of sensors are there, however each sensor can sense only specific criteria (temperature, pressure, sound etc.) and this information is maintained in the cluster head. Therefore upon receiving the query, cluster head sends the particular query to those limited number of sensors which can answer that particular query. Hence the communication cost during the execution phase is likely to be reduced to the greater extent.

V. Conclusion

In this paper, we purposed the model for selection of cluster heads depending upon their working condition and energy level. This ensures longer life for each cluster head, since sensor with maximum energy is selected as new cluster head. This reduces the setup cost in the initial phase of the query.

In the query execution phase, communication cost is expected to be reduced since the query is being sent only to the nodes which are meeting the required criteria, as detailed above. In future work, this model and the algorithm will be implemented using MATLABsimulator. Further comparisons will also be carried out with other models for the same purpose given in the literature.

References

[1]. Ousmane Diallo, Joel J. P. C. Rodrigues, Senior Member, IEEE, Mbaye Sene, and Jaime Lloret, Senior Member –"Distributed Database Management Techniques for Wireless Sensor Networks",*IEEE Transcations on Parallel and Distributed Systems,IEEE 2013*

[2]. MeenakshiDiwakar, Sushil Kumar – " An Energy Efficient Level Based Clustering Routing Protocol For Wireless Sensor Networks", *International Journal Of Advanced Sensor Network Systems, Vol 2, No 2, April 2012*

[3]. Devesh Pratap Singh, R. H. Goudar , Bhasker Pant , Sreenivasa Rao-" Cluster head selection by randomness with data recovery in WSN" , *Aug 2014, Springer*

[4]. C. Komalavalli, Chetna Laroiya Energy Efficient Protocol in Wireless Sensor Network: Reactive On-Demand Routing, *IITM Journal for Management & IT, Vol 2 Issue 1,2010*

[5]. SrieVidhya Janani. E, Ganeshkumar.P, Vasantha Suganthi.G, Sultan.M, Kaleeswaran. D," A Survey on Algorithms for Cluster Head Selection in WSN", *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 2, No 5, May 2013*

[6]. S. Bandyopadhyay, E. Coyle, An energy efficient hierarchical clustering algorithm for Wireless sensor networks, in: *Proceedings of the 22nd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2003), San Francisco, California, 2003*

[7]. W. Heinzelman, A. Chandrakasan and H. Balakrishnan, "Energy- Efficient Communication Protocol for Wireless Micro sensor Networks" *Proceedings of the 33rd Hawaii InternationalConference on System Sciences (HICSS '00), 2000, pp. Vol.2:10.*

[8]. S. Lindsey, C. Raghavendra, "PEGASIS: power-efficient gathering in sensor information Systems," in: *Proceedings of 2002 IEEE Aerospace Conference, Big Sky, Montana, USA,vol.3, 2002, pp.1125-1130*.

[9]. A. Manjeshwar, D. Agrawal, "TEEN: a routing protocol for enhanced efficiency in wireless sensor networks," in: *Proceedings of 15th International Parallel and Distributed Processing Symposium (IPDPS'01), San Francisco, CA, 2001, pp. 2009-2015.*

[10]. O. Younis and S. Fahmy, "HEED: "A Hybrid Energy-Efficient Distributed ClusteringApproach for Ad Hoc Sensor Networks," *IEEE Transactions on Mobile Computing, 2004,Vol. 3, No. 4.*