Exploration of Data Behavior at Mobile Sink

Sumedha Sirsikar
Department of Information Technology
Maharashtra Institute of Technology
Pune, India
sumedha.sirsikar@mitpune.edu.in

Shalini Gour
Department of Computer Engineering
Maharashtra Institute of Technology
Pune, India
shalini.gour@mitpune.edu.in

***Abstract*— Wireless Sensor Networks (WSN) consist of small nodes with sensing, computation, and wireless communications capabilities. These sensors have the ability to communicate either among each other or directly to an external base-station (BS). Sensor nodes are highly constrained in terms of energy supply and bandwidth. Wireless sensor network with static sink suffers from the Energy Hole Problem [2] or Hot Spot Problem. In this problem nodes near to the sink depletes more energy which leads to network partitioning. We propose a solution to this problem as mobile sink in wireless sensor network. Using mobile sink the problem of Hot Spot can be efficiently minimized. The proposed protocol uses delay tolerant approach where the proposed algorithm divides the WSN into the number of clusters. Each cluster has a cluster head and cluster members when a node wants to send data to the sink it first sends data to the cluster head (CH) then CH sends data to the sink. In the proposed work the network is scalable so if the network size increases the sink will store a huge amount of data such as location of CH and aggregated data from various CH. Hence the proposed approach also incorporates the behavior of data at the mobile sink.**

Keywords— Wireless Sensor Network, Clustering, data management

I. Introduction

Wireless sensor networks are the network of the sensor nodes. Sensor nodes are the battery powered devices, so to increase the lifetime of the network efficient management of energy is essential. In WSN, sensor nodes are deployed randomly or sparsely. These sensor nodes need to communicate to the sink. If the position of the sink is static then that WSN is said to be wireless sensor network with static sink. In wireless sensor network with static sink a sensor node forwards the data in a multi-hop manner to the base station. In this case the node which is near to the sink depletes more energy as compare to the other nodes in the network. So the nodes near to the sink dies early and causes network partitioning, this problem is known as Energy Hole Problem [8] or Hot Spot Problem [8].

Apart from Hot Spot solution the mobile sink has many advantages such as load balancing, security, shorter data dissemination path and better handling of sparse or disconnected network. Frequent change of neighboring node of the sink leads to balance the load of the network. Shorter data dissemination path provides longer network lifetime by increasing throughput and decreasing energy consumption. Other advantage of the mobile sink includes security benefits. In wireless sensor network with mobile sink an adversary has to find and trace the location of the sink to attack on the data stored on the mobile sink which is not easy. Besides the energy hole problem the sensor network faces another challenge of data management at the sink. When the network is a distributed dense wireless sensor network. There may be millions of sensors and data generated by millions of sensor is very large in volume. In the proposed work the network is a distributed dense wireless sensor network. In case of dense and scalable network the data may be generated from millions of sensor at a time. Millions of sensor nodes may generate bigdata.

The challenge here is to explore the behavior of data so that the single mobile sink can propagate data to the base station reliably and accurately. Another challenge is limited energy (i.e. Hot Spot Problem) of the sensor because if the sensor dissipates more energy it will die early and thereby reducing the network lifetime.

The solution to Hot Spot problem is to make the sink mobile. Wireless sensor network that has a mobile sink is known as wireless sensor network with mobile sink. The network with mobile sink reduces the possibility of network partitioning. The sink moves around the network to collect the data from the various sensor nodes. This movement usually changes the neighboring node of the sink. Hence it minimizes the problem of Hot Spot. An example of wireless sensor network with mobile sink is traffic surveillance. In traffic surveillance application the sensor nodes are deployed homogeneously along the highway. These sensor nodes collect the data about the car which is speeding along the highway. The sink is equipped in a car of the police. This police car is driven along the highway once in a day to collect data (i.e. pictures) from the sensor node. Other application of mobile sink includes habitat monitoring, fire detection system, battlefield surveillance, smart house and hospitals and pollution control etc. [2].

In this paper we propose Exploration of data behavior at the sink in a clustered wireless sensor network. Since the data received by the sink is a sensed data from millions of data. To manage this much amount of data by a single sink is a huge challenge. As a solution to this problem the proposed work explores the behavior of data in terms of gathering and delivery of data.

In the proposed approach initially all sensor nodes are deployed randomly and uniformly. The wireless sensor network is organized in the form of the cluster. In the proposed algorithm the clusters are formed only once during the lifetime of the sensor network. The proposed approach makes use of single node to single sink communication. The CH assigns TDMA schedule to all cluster members for collecting data from all its members. The CH aggregates data from all of its member and after aggregating it delivers data to the sink. The proposed approach is a delay tolerant approach. In this approach the CH waits for the most favorable position of the sink and when sink is in the correct position the CH sends the data to the sink. The proposed approach provides better network lifetime by decreasing energy consumption and increasing throughput. Along with solving the hot spot problem the algorithm also focuses on the data behavior at the mobile sink. In the proposed work the sink maintains the record of the location and residual power of CH. Sink also gathers data sent by different CHs. since every CH delivers the aggregated data to the sink so the volume of data at the sink is large. This way data gathered by the sink becomes big data. The rest of the paper is organized as follows: Section II describes the survey of existing algorithm. Section III explains the proposed system. Implementation details i.e. algorithm and flowcharts are given in section IV. Concluding remarks and direction for future work have been included in section V.

II. Related Work

We propose a comprehensive review of existing protocols and approach and our goal is to provide the state-of-the art in the subject by including the most recent approaches and protocols. Here we discussed most of the related work for cluster based wireless sensor network with mobile sink.

Zhao et al [1] the proposed approach uses a clustered wireless sensor network and multi-hop inter-cluster routing to send data to the sink. According to the proposed routing algorithm the network is first divided into clusters than each cluster head first collects data from there corresponding cluster member and then form a inter cluster routing to transfer data to the sink. Establishment of the inter-cluster routing path is controlled by the sink. The update in the inter-cluster routing path is notified by the sink. Sink compares its displacement with the threshold of distance whenever it needs to update the inter-cluster routing path. The author proposed DMS-CTLR and RMS-CTLR. The DMS-CTLR is based on Displacement update of two levels routing among the cluster head with the mobile sink and RMS-CTLR is based on round update CTLR (two levels of routing among the cluster head) with mobile sink. And with the help of simulation results they proved that DMS-CTLR consumes less energy. Nazir et. al. [2]. The aim of the proposed approach is to mitigate with the Hot Spot Problem. Each cluster head send their data to the sink when sink comes in its vicinity. The sink moves towards the CH which is having highest residual energy. The proposed approach works in two cycles in first cycle all Ch gets registered with the sink. In second cycle sink is having all CH(s) with their residual energy. When sink has to move to any of the CH it sorts all CH(s) according to their residual energy and moves to the CH which is having highest energy. In this way mobility of the sink is controlled. Suganthi et. al.[3] $:-$}This approach makes use of multiple sink to transfer data to the base station. The whole network is divided into number of clusters. When an event is created by the sensor node it first sends data to the corresponding CH. CH aggregates that data received from that sensor node with other sensor node (cluster member) of the cluster. After aggregating sink checks for the availability of any of the three sink moving in circular, rectangular and linear motion respectively. If any of the sink is available it sends data immediately if no sink is available then it forwards data to the next neighboring CH. Similar to the energy saving cluster routing it does inter-cluster routing to send data but in this approach it always tries to find the shortest path using a shortest path algorithm. In this approach event is created by both sensor node and sink. Sink periodically sends its interest in receiving data and also when sensor node wants to send data to the sink it can also initiate.

Shekastehband et. al.[4] In this paper a hierarchical wireless sensor network with mobile sink is proposed with best mobility pattern so that to have an energy efficient data communication between the sink and the CH. A special PSO based algorithm is used to find the best mobility pattern. In the proposed approach the sink moves to the new location uploads the data from the CH. The network lifetime is split into number of rounds in each round the sink relocates itself and the update in the location is notified to the CH so that minimum energy cost is incurred in data transmission. Wang et. al.[5]. The proposed approach is an energy efficient approach. The network is divided into clusters. There is intra-cluster routing to send data to the CH. The sink moves around the edges of the network and the mobility pattern of the sink is fixed and predictable. Similar to PSO based approach discussed above; selection of CH is based on the residual energy. Whenever a node wants to send data to sink it first sends data to its corresponding CH. This sending of data to the CH can be single hop and can be multi-hop depending up on requirements. The node which is sending data to the CH is very far it can send data using multi-hop strategy. When CH aggregates data of its entire cluster member it sends this aggregated data to the sink when sink is available.

III. Proposed System

This is a delay tolerant application. According to the proposed algorithm the sensor network is divided into number of clusters, each cluster has a corresponding cluster head. Member of each cluster communicates to their respective cluster head when there is data to be transmitted. For the purpose of cluster formation the FLOC (Fast Local Clustering Service) [9] protocol will be used. Using FLOC protocol a non-overlapped and energy efficient clusters will be formed. Once cluster formation is done each cluster head send their location and ID to the sink using PEGASIS [10] chain based routing protocol. This protocol is used to pass the location notification to the sink. Sink maintains a table which stores the address of all the cluster head. The sink is using a data driven approach. Sink periodically moves around the network territory and when it comes near to one of the cluster head it sends a beacon packet to the nearest cluster head which receives this beacon packet transmit their data packet to the sink. The proposed approach consumes less energy and incurs less delay.

The proposed system is divided into four phases: First phase is cluster formation phase second phase is CH location sending phase, third phase is mobility management and data transmission phase and fourth phase is Analysis of behavior of data at the sink.

A. Cluster Formation Phase:

In the cluster formation phase, the process of cluster formation is done. Here To form good and efficient clusters we are using FLOC protocol. This protocol provides an algorithm according to which the whole network is divided into number of equal sized, non-overlapped clusters. A node can communicate reliably with the nodes that are in its inner-band (i-band) range, and unreliably (i.e., only a percentage of messages go through) with the nodes in its outer-band (o-band) range.

According to this protocol all the nodes which are at the unit distance from the cluster head are known as I-Band node and all those nodes which are far away(more than two units away) are known as O-Band node. The FLOC works in way so that a non-overlapped clusters are formed. According to this protocol all the node maintains two variable $Cluster_{id}$ and $CH_{status.}$

$Cluster_{id}$ stores the ID of the cluster to which non-cluster member belongs.

CH_{status} { J.Idle, J.Cand, $J.C_{head}$, I-Band ,O-Band}

J.Idle: when j is not part of any cluster.

j.Cand: when J wants to be a cluster Head

$J.C_{head}$: when J is cluster Head

J.I-Band: J is an inner-Band member of a cluster head.

j.O-Band: J is an outer-Band member of the cluster head
Whenever a node J wants to become first it will compute its ch_{prob} using the formula given by

$$Ch_{Prob} = c_{prob} * E_{res} / E_{max} \quad (1)$$

Where:
c_{prob}= initial percentage of cluster heads among all n nodes
E_{res}= estimated current residual energy in the node
E_{max}= Maximum energy of a node

Firstly a node which is ideal and wants to become cluster head will calculate CH_{prob}, if Ch_{prob} of this node is equal to one than it will discover all its I-Band member and broadcast a candidate message to all other nodes. If any of the recipient node discovers that the recipient node is also in I-Band of new Candidate it will broadcast a conflict message. If the new node receives a conflict message it will give up its candidacy and will update its status as an O-Band member of sender's CH (means it will make its $cluster_{id}$ of the Cluster Head which sent conflict message to it) otherwise if it does not receive any conflict message than it will become cluster head and will broadcast $Ch_{message}$ to all other nodes in the network. The other entire nodes which receive this message will update their I-band or O-band status accordingly for all the nodes whose probability is not equal to one or less than one will be considered as O-band member. In future they can update their status from O-Band to I-Band for some other cluster head.

B. CH location Sending Phase:

After formation of clusters all the cluster head send their location and ID to the sink. For this purpose we will use PEGASIS protocol.

PEGASIS Protocol Works in two Phases: PEGASIS (Power-Efficient gathering in Sensor Information Systems), which is a near optimal protocol for high rate data gathering applications in sensor networks. The key idea of the PEGASIS protocol is the formation of a chain among the sensor nodes so that each node will receive from and transmit to a close neighbor. Gathered data moves from node to node, get fused, and eventually a designated node transmits it to the BS. To send location of CH to the sink, PEGASIS first will form chain with the help of neighboring CH then it elects a leader for each chain. This leader will collect the location of all CH coming along its own chain and sends the location to the sink.

Chain Formation by Cluster Heads: PEGASIS is a protocol which is basically used for data propagation. The chain formation phase of pegasis is as follows: The chain formation phase starts from the farthest CH from the sink. The farthest CH selects the CH which is nearest to this CH as the next CH in the chain formation process. If it finds the nearest CH it computes following three parameters: d_{avr} , d_p and d_{comp}. then it checks whether the distance of new CH is less than the d_{comp} if the distance of new CH is less than the d_{comp} than it will check if the CH already exist in the chain or not if the CH already does not exist in the chain than it adds this CH to the chain otherwise it searches for another CH which is having less distance or which already does not part of any chain.

Sending Location of the CH to the Sink: In this phase of pegasis, location of the CH is gathered from the neighboring node and then transferred to the sink. Each chain has a leader which takes responsibility to transfer the location of CH to the sink. The selection of leader is done randomly using the formula $CHAIN_{member}$ MOD N the location of the chain leader is also considered as random. This leader of the chain uses a CONTROL TOKEN PASSING APPROACH to transfer location of CH to the sink. In this approach the leader of the chain sends a control token to the node which is prior to this leader they send data to their nearest next CH that next CH receives data from the previous one fuses its own data and transmit this data to the leader.

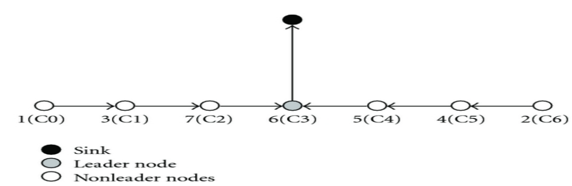

Figure 1: Sending location to the Sink

As shown in the Figure 1, C3 is the leader of this chain, this leader will send a control token to the CH, C0. When C0 receives the token it will send its data to the next CH in the chain that is C1, than C1 collects data from node number C0, fuses its own data and send it to node number C2, than C2 fuses its own data and sends it to the leader CH after this leader passes the token to C6 the same process is repeated in this side and when all data is collected by leader CH. The leader sends all collected location to the sink.

C. Mobility Management and Data Transmission Phase:

In this phase when all clusters are formed and sink knows the location of each cluster head so it will move around the network based on the location information. It stays at the cluster head location for some specific time to collect the data. When the sink wants to collect the data from the CH, it will send a beacon packet to the CH. The CH starts transmitting the data to the sink. If sink does not receive any data within the specified time limit it will move from that location. In that way the mobile sink will collect the data and move in to the network. Instead of random movement we make the mobile sink to move in a predefine path. So that It can cover the network uniformly.

D. Analysis of Behavior of Data at the Sink:

In the proposed work the sink stores the location of each CH and the data sent by the CH. Sometimes (when the network size increases) the data collected by the sink may be large in volume as the sink collects aggegated data from various CH. In this phase an intensive analysis is done on the data stored by the sink. This analysis includes the behavior of data in terms of data gathering and delivery of data.when the sink gets all the information of CH it will find how many CH are in the coverage with respect to the current position of the sink. Once it will find some no CH then it will find which CH is having least residual energy. Then the sink will send the beacon packet to the CH which is having least residual energy. Then again sink will see which CH is having next least residual energy. In this way it will cover all the CH which are in the coverage of current position of the sink. At the time of processing the data, the sink checks the correctness of the data.

Following are the steps of algorithm:

1. Once the clusters have been created, each CH will send its ID and its residual power to the sink.

E_res= Total_Energy-Consumed_Energy (2)

Energy Consumption=Initial Energy-Remaining Energy (3)

2. Sink will store the info of various CH and its residual power. Then it will find how many CHs are in the coverage in the current position of the sink.

3. When it finds number of CHs are in the coverage then it will set a minimum threshold (Assuming 10 joule) of energy then all the CH which comes under this threshold it will send beacon packet to only those CHs. In this way at a time the sink will handle a subset of data.

The proposed approach concentrates on the behavior of data so that there should be proper data management at the sink.

Advantages of the proposed system:

- Throughput has increased
- Better Network Lifetime
- Consumes less energy
- Imposes a limit on formation of number of clusters
- Efficiently mitigates the problem of Hot Spot
- Mobility Pattern is controlled therefore less latency in data delivery

IV. IMPLIMENTATION DETAIL

Algorithm for cluster formation phase:

```
Cluster Formation
Step1:S_Node = { Set of all nodes which wants to
      become CH }
Step 2: Compute: Ch_Prob = C_Prob * E_Res / E_Max
Step3:  If( Ch_Prob == 1)
Step 4: {
Step 4: Discover all the nodes which are in I-Band
      and O-Band
Step 5: BCAST(Cand_Message)
Step 6:  if ( j.I.Band belongs to i.I-Band)
Step 7 : {
Step8 :       BCAST(Conflict_Message)
Step9 :  J.Status = O-Band;
Step10: J.Cluster_ID = I;
Step 11:  }
Step 12:  else
Step 13:  {
Step14:     J.Status = C_Head;
Step15:     BCAST(CH_Message);
Step 16:    }
Step 17:    }
Step 18: else if(  j. Rcv.msg <- C_head_msg_i)
Step19:  {
Step 20:  j.Status = I-Band_i / O-Band_i
Step 21:   j.Cluster_id = i;
Step 22:   }
Step 23:  else
Step 24:  {
Step 25: if ( J == O-Band node && J belongsto I-Band_I)
Step26: {
Step 27:     J.status = I-band;
Step28:      J.Cluster_ID = I;
Step 29:  }
Step 30:  }
```

Flow Chart for Cluster Formation phase:

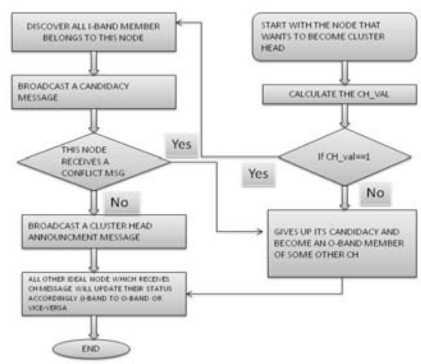

Figure 2: Flow Chart for cluster formation

Flow Chart for Location Sending Phase is shown in Figure 3.

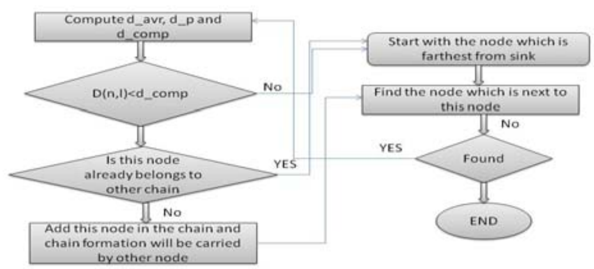

Figure 3: Flowchart for forming chain

V. Conclusion and Future Work

The challenging issue related to WSN is how to utilize limited energy of the sensor node so that better network lifetime can be achieved. In this paper we intuitively focused on this issue and came up with a solution in which we have made sink as a moving entity. Making the sink moving efficiently mitigates the energy hole problem. Proposed clustering algorithm minimizes the energy consumption in data transmission. A uniform formation of clusters is a contribution in increasing the throughput. The proposed work also focused on the data behavior at the sink. The behavior of the data is analyzed in terms of gathering and delivery of data. The proposed algorithm solves the problem of handling big data up to much extent.

As in our work we did not focus on the situation when CH itself dies. Therefore in our future implementation we will include a method or possibly another phase in which a new CH will be created if any of the current CH dies. This selection of new CH will be based on highest residual energy of the sensor.

Refeerences

[1] Chenglin Zhao, Yongxing Wang, Xuebin Sun and Ting Jiang, "An Energy Saving Cluster Routing for Wireless Sensor Network with Mobile Sink" , International Conference on Advanced Intelligence and Awareness Internet, pp 113-117,2010.

[2] Babar Nazir and Halabi Hasbullah, "Mobile Sink Based Routing Protocol for Prolonging Network Lifetime in Clustered Wireless Sensor Network (MSRP)", International Conference on Computer Applications and Industrial Electronics (ICCAIE) pp 624-629,2010.

[3] K.Suganthi , Dr. B.Vinayaga Sundaram ,K.S.Vishwa Kumar, J.Syed Ashim and S.Saravana Kumar, "Improving Energy Efficiency and Reliability using Multiple Sinks and Hierarchical Clustering In Wireless

[4] Sensor Network with Mobile Sink",International Conference on Recent Trends in Information Technology (ICRTIT),pp 257-262,2011.

[5] S. Shekastehband, M.H. Yektaie and M. Abbasi Dezfouli, "PSO Based Localization Strategy in Wireless Sensor Network with Mobile Sink", 2nd International Conference on Control, Instrumentation and Automation (ICCIA), pp 815-818, 2011.

[5] Jin Wang, Yue Yin, Jeong-Uk Kim, Sungyoung Lee and Chin-Feng Lai, "A Mobile Sink Based Energy Efficient Clustering Algorithm for Wireless Sensor Network", IEEE 12th International Conference on Computer and Information Technology (CIT),pp 678-683, October-2012

[6] C. Chen, J. Ma, and K. Yu, "Designing Energy Efficient Wireless Sensor Networks with Mobile Sinks.", Proc. of ACM Sensys'06 Workshop WSW, Boulder, Colorado, 2006.

[7] F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci "A survey on sensor networks.", IEEE Communications Magazine, 40(8), pp. 102–114, 2002.

[8] Tunca, Can, Sinan Isik, M. Donmez, and Cem Ersoy. "Distributed MobileSink Routing for Wireless Sensor Networks: A Survey." pp1-21, May-2013.

[9] Murat Demirbas, Anish Arora, Vineet Mittal," FLOC: A fast local clustering service for wireless sensor networks "Workshop on Dependability Issues in Wireless Ad Hoc Networks and Sensor Networks (DIWANS/DSN), 2004.

[10] Lindsey, S. Raghavendra,"PEGASIS: Power-efficient gathering in sensor information systems", IEEE Aerospace Conference, pp n1125-1130, 2002.

[11] K. I. Hwang, J. In, and D. S. Eom, "Distributed dynamic shared tree for minimum energy data aggregation of multiple mobile sinks in wireless sensor networks", Proceedings of the 3rd European conference on Wireless Sensor Networks, Springer, Berlin, Germany, vol. 3868, pp. 132-147, 2006

[12] Suraj Sharma and Sanjay Kumar Jena, "Data Dissemination Protocol for Mobile Sink in Wireless Sensor Networks", Journal of Computational Engineering 2014 (2014), Article ID 560675, pp 1- 10, 2014.