Hierarchical Clustering in Decentralized Lifetime Maximizing Tree for Data Delivery in Wireless Sensor Networks

Ankita Bansal
University School of Information Technology
M.Tech CSE, GGSIPU, Dwarka, New Delhi, India

Jyotsna Singh
University School of Information Technology
GGSIPU, Dwarka
New Delhi, India

ABSTRACT
A wireless sensor network (WSNs) is a self configuring network which is an anthology of thousands of sensors having insufficient storage, battery life and computation capabilities that communicate via wireless fashion. Wireless sensor network have diverse application domain which includes habitat monitoring, surveillance etc. For data aggregation at single source antecedent to transmitting to ay distant user, decentralized maximizing tree is created that preserves energy and maximizes lifetime of event sources. In this paper we propose a decentralized lifetime maximizing tree along with hierarchical clustering that will minimize the energy consumption, delay during data collection and reduce time complexity.

Keywords
Lifetime Tree, Hierarchical Clustering, Wireless Sensor Network

1. INTRODUCTION
Wireless sensor networks are becoming very popular that comprise of large number of miniature sensor nodes equipped with limited processing and battery power communicating among themselves over well defined protocols using radio signals, and deployed in quantity to sense, monitor and understand the physical world. Since sensors have less battery life therefore data aggregation limits the lifetime for which sensor remains active. The limited battery leads to disruption in the connections from the network. These disruptions suggest that the design must incorporate topological changes. Since the dense deployment of sensors nodes leads to the detection and transmission of data from the nearby nodes upon receiving a single stimulus. Thus, idea is not to allow the direct transmission of to interested users upon event detection instead aggregating them to remove redundancy [10]. The application domain of WSN is still expanding therefore it is important to support the data aggregation scheme from multiple nodes for simultaneous and fast processing of the data.

In this paper, we focus on construction of decentralized lifetime maximizing tree using energy efficient hierarchical clustering in which along with the tree of nodes created inside the clusters hierarchy of clusters is also maintained to reduce delay and time complexity and minimize energy consumption.

This idea consists of three parts, namely, clustering of node using energy efficient hierarchical clustering and creating hierarchy[1], construction of decentralized lifetime maximizing tree within each cluster [8], and aggregating the data collected from the WSN nodes by applying a cluster scheduling approach to transfer it , which uses HyMac[7] mechanism.

The rest of the paper is organized as follows. Section 2 describes some of the related works and reasons for not using them. In section 3 we illustrate our approach. Section 4 has the proposed algorithm. Finally, concluding remarks are provided in Section 5.

2. RELATED WORKS
In last few years many studies have achieved data aggregation using various schemes, namely, mobile sink, LEACH (Low-Energy Adaptive Clustering Hierarchy) [4], Directed Diffusion [2] and Max-Min d-Cluster Algorithm [5]. All these schemes have tried to enhance network lifetime and lessen the energy consumption. In mobile sink approach network lifetime is raised by four times as compare to the network in which sink is static but it suffers from serious faults as there is an increased physical delay due to which the physical mobility of the sink is reduced in the wireless communication. It depletes the battery life unnecessarily. LEACH is a self-organizing, probabilistic, one hop and adaptive clustering protocol. LEACH is used to improve lifetime of WSNs by trying to evenly distribute the energy consumption among all the nodes of network and to reduce the energy consumption in network nodes by performing data aggregation. It forms clusters based on received signal strength and also uses the cluster head nodes as routers to the BS. Data aggregation is local to the clusters. Each cluster has a cluster head, which communicates with every node of that cluster. The sink aggregates data, transmitted by cluster heads, from other nodes. Since a cluster head loses energy due to repeated transmissions, the cluster head is re-selected. The decision on cluster head election and rotation is probabilistic therefore a very low energy node can be selected as cluster head due to which good cluster heads distribution cannot be guaranteed. Directed Diffusion involves two types of messages, namely, the "Interest" message and the actual data messages.

To aggregate data by using Directed Diffusion, sink node requests data by sending "interest" message for named data that consists of a time-to-live value, and also the addresses of the source and destination nodes. When destination node in that region receives an "interest" message, it activates its sensors which start collecting information about pedestrians. When the sensors report the presence of pedestrians, this information returns along the reverse path of interest propagation. Intermediate nodes might aggregate the data by combining reports from several sensors. If the downstream nodes cannot be reached by the "interest" message from the current source then the current destination becomes the source node by changes its address, reduces the time–to-live value and rebroadcasts the "Interest "message. The Max-Min d-Cluster Algorithm generates d-hop clusters with a run-time of O (d) rounds. But this algorithm does not ensure that the energy used in communicating information to the information centre is minimized.

3. PROPOSED APPROACH

We have discussed that our main objective is to reduce energy consumption, delay, and time complexity and maximize network lifetime and efficient utilization of bandwidth. To resolve the problem we partition the sensor network into clusters using energy efficient hierarchical clustering [1] based on probability of being cluster head and number of hops from cluster heads and then arranging clusters into different hierarchy levels. Then a decentralized life maximizing tree [8] is constructed within the cluster choosing a parent closest to the sink node to serve as sub- sink. Once the sub sink is chosen, scheduling of cluster is done using FDMA approach in which we allocate a frequency channel to each sub sink from the existing pool. If frequency channel is not available then we allocate half frequency to a sub-sink from a low data rate transferring cluster. When the channel is free it is reallocated to another cluster. The sink will broadcast a topology packet containing information of the network as which source nodes are attached to which sub- sink node as per their location [3]. The hybrid TDMA/FDMA channel access technique [7] is used according to which sink node broadcasts a schedule packet that inform other nodes about their time slots as well as their channel frequencies for exchanging messages. According to [10] single sleep awake concept is used in which the source nodes wake up only once to listen and to transmit and rest of the time, they will remain in sleep state. We include a concept of LPL (low power listening) [6] the nodes are in LPL state all the time to gauge topology changes and if there is a topology packet coming their way they wake up and make necessary changes.

We revise Decentralized lifetime maximizing tree construction algorithm [8] to save the functional lifetime of all sensor nodes and efficiently use the energy of the source nodes. In DLMT [8] tree is created in which node with maximum residual energy is choose to be the parent node that act as hub for data aggregation. According to this algorithm all nodes are arranged in such a way that each parent has maximal existing resources so that it can receive data from all of its extends the time to refresh the tree and minimize the amount of data lost due to a broken tree link before the tree reconstructs. In the proposed method follow the approach of hierarchical clustering [1] of nodes based on node probability of becoming cluster head and number of hops.

Using this algorithm multi hop clusters are developed and arranged in hierarchy. By maintaining this hierarchy network lifetime will be maximized and delay and time complexity will be minimized. The clusters formed using this algorithm goes through algorithm called Decentralized Lifetime Maximizing Tree using Hierarchical Clustering (DMLTHC) which create trees within the clusters already created and arrange them into a hierarchy. The choice of the tree is based on the minimum distance of the sub-sink from the sink.

At the end we apply HyMac [7] algorithm which the communication period is a fixed-length TDMA cycle composed of a number of frames. Each frame is equivalently divided into several fixed time slots where slot duration is the time required to transmit a maximum sized packet.

4. PROPOSED ALGORITHM

4.1 Energy Efficient Hierarchical Clustering Algorithm

Energy Efficient Hierarchical Clustering algorithm (EEHCD) [1] is a k-hop, distributed hierarchical clustering algorithm in which each sensor is elected as a cluster head with probability "p_i" and broadcast its selection to the neighboring nodes within its communication range and also known as "volunteer" cluster heads. After this all the sensors within "k_i-hop" distance from a "volunteer" cluster head should receive an election message either directly or through forwarding by in between nodes. When any sensor node receives that message and is not itself a cluster head, becomes a member of the closest cluster. Nodes that are neither cluster heads nor belong to a cluster are forced to become cluster heads. Specifically, if the election messages do not reach a node within a preset time interval t, the node becomes a "forced" cluster head. There are "h" levels in the clustering hierarchy with level 1 being the lowest level and level h being the highest.

This algorithm groups the sensor nodes into clusters based on k and p parameter and hence ensure minimization of data transfer delays and time complexity and maximization of lifetime.

EEHCD Algorithm

E[C]: expected total cost of communicating information from sensors to the processing centre.

λ: Poisson process of intensity

N: sensors in a square area of side 2a is a Poisson random variable with mean A*λ

$A = 4a^2$: is the area of the square

n: number of sensors in a given square area

r: radio range

E[C] = E [E[C | N=n]]

$$= \lambda * A * \prod_{i=1}^{h} \text{Ceil}\left[(pi * 0.765 * a)/r\right] + \lambda * A * \sum_{i=1}^{h} (1-p_i) * \prod_{j=1}^{i-1} (p_j) * \text{Ceil}\left[1/\left[2r * \left(\lambda * \prod_{j=1}^{i} p_j\right)^{1/2}\right]\right]$$

4.2 Decentralized Lifetime Maximizing Tree with Hierarchical Clustering Algorithm

Now we apply DLMT [8] through which we create a tree of sensor nodes within each cluster which are already arranged into different hierarchical levels.

The tree construction leads to minimization of delay and energy consumption. Now we compare the trees based on their distance from sink chooses the one with minimum distance from the sink. The final choice of the tree is done by taking into account the distance of the sub-sink or parent from the sink node.

BestDLMTC (DLMTHCi, DLMTHCj, di, dj)

1. if rows in (DLMTHCj) > rows in (DLMTHCi)
2. return true
3. if [rows in (DLMTHCj) = rows in (DLMTHCi)] and [DMLTEj > DLMTHC_Ei]
4. return true
5. if [rows in (DLMTHCj) = rows in (DLMTHCi)]and [DMTLEj = DLMTC_Ei] and [tree depth of DLMTHCj < tree depth of DLMTHCi]
6. return true
7. if [rows in (DLMTHCj) = rows in (DLMTHCi)] and [DMLTEj = DLMTHC_Ei] and [tree depth of DLMTHCj = tree depth of [DLMTHCi] and [ej > ei] and [dj < di]
8. return true
9. if [rows in(DLMTHCj) = rows in(DLMTHCi)] and [DMLTEj = DLMTHC_Ei] and [tree depth of DLMTHCj = tree depth of DLMTHCi] and [ej = ei] and j < i
10. Return true Else if return false.

4.3 Applying FDMA_SINK [10] on the DLMTHC Tree

We partition all the available frequencies in the band into “R” frequency ranges based on the number of clusters “K. Primarily all the sub-sinks send a request for allotment of the frequencies.

Sink then checks for the whether the frequency range is available if yes it will be assigned. Otherwise sink may withdraw a frequency from some other cluster if its data transmission rate is low and assign that to some cluster.

SSi : sub-sink of ith cluster

Fi : frequency band for ith cluster

FDMA_SINK (DLMTHC Treei, Ki, Fi, Ci)

1. Calculate the range Ri = Fi/Ki
2. For each S Si ϵ Ci where i=1 to k
3. Assign Ri
4. If SSi receives the last packet then

   Withdraw the frequency and add it to the free pool.
5. End for
6. Request for new Fi allotment
7. If available from free pool, Assign from free pool.
8. Else withdraws half the frequency range from a low data rate transfer cluster and assign to the subsink.

4.4 Cluster Scheduling Using HyMac

We apply hymac [7] scheduling algorithm on the DLMTHC tree having the base node as its root. A default time slot and a frequency via FDMA_SINK () function is assigned to each node Ni traversed by DLMTHC. Then the possibility of having an interference with any of its same-height previously-visited one-hop AND two-hop neighbors is checked. If a conflicting neighbor Ni is found for Ni, the algorithm checks whether Ni and Nj are siblings. If so, Ni will be assigned a different time slot than that of Nj. If they are not siblings then Ni will be assigned a different frequency than that of Nj, allowing both Ni and Nj to send messages to their parents at the same time slot but in different channels.

When DLMTHC is about to start a new level (height) of nodes the default time slot number will be increased by one.

Once all nodes are processed according to the above heuristic, the entire time slot assignments will be inverted such that the slot number assigned to every node is smaller than that of its parent. This inversion is done as following:

$$t_{new} = t_{max} - t_{current} + 1$$

Where tnew is the new inverted assigned slot, tcurrent is the current slot number assigned to the node and tmax is the total number of assigned slots. Note that such an assignment allows the data packets to be aggregated and propagated in a cascading manner to the base station in a single TDMA cycle.

Require: A Graph of Sensor Network Topology

Ensure: An scheduled Tree of the given network

```
HyMAC with clusters Scheduling Algorithm
ENQUEUE (Q, S)
while Q is not empty do
  v ← DEQUEUE (Q)
  timeSlot[v] ← currentTimeSlot
  FDMA_SINK () /* assigning channels */
  for all Visited same-height 1-2-hop nbr n of v do
    if parent[n] == parent[v] or #Channel >= available chnls then
      if timeSlot[v] = timeSlot[n] then
        timeSlot[v] ← timeSlot[n] + 1
      end if
    else
      If timeSlot[v] =timeslot[n] and channel[v]=channel[n] then
        channel[v] ← channel[n] + 1
      end if
    end if
  end for
```

```
  for all unexplored edge e of v do
    let w be the other unvisited endpoint of edge e
    parent[w] ← v
    height[v] ← height[w] + 1
  end for
end while
```

The overall complete process is:

Final Algorithm

Require: Set of sensor nodes

Step 1: Apply EEHCD () /* Creates clusters and maintain hierarchy */

Step 2: Creating decentralized lifetime Maximizing tree using DLMTHC ().

Step 3: Choosing the life maximizing tree using Best DMLTHC ().

Step 4: Scheduling Using HYMAC () which calls FDMA_SINK for scheduling based on time slots and frequency range.

5. CONCLUSION

In this paper we propose the Decentralized Lifetime Maximizing Tree with hierarchical clustering construction algorithm. Clustering done on the basis of distance ensures close proximity of the nodes and hierarchy of clusters is maintained which ensures more reduction in data transfer delays. Energy conservation is done by waking nodes once instead of twice using sleep awake concept leads [14] to further reduction in data transfer delays, thus utilizing the energy available effectively. Bandwidth is efficiently utilized through allotment of frequencies from the free pool or withdrawing half frequency from cluster with low data transfer and assigning it to some other cluster using HyMAC [7] technique.

6. REFRENCES

[1] “An Energy Efficient Hierarchical Clustering Algorithm for Wireless Sensor Networks” By Seema Bandyopadhyay and Edward J. Coyle School of Electrical and Computer Engineering Purdue University West Lafayette, IN, USA.

[2] “Directed Diffusion for Wireless Sensor Networking” By Chalermek Intanagonwiwat, Ramesh Govindan, Deborah Estrin, John Heidemann,and Fabio Silva.

[3] “PEDAMACS: Power Efficient and Delay Aware Medium Access Protocol for Sensor Network” By Sinem Coleri Ergen, IEEE, 2006.

[4] “Energy Efficient Communication Protocol for Wireless Sensor Network” By Wendi Rabiner Heinzelman, Anantha Chandrakasan and Hari Balakrishnan Massachusetts Institute of Technology Cambridge, MA 02139.

[5] “Max-Min D-Cluster Formation in Wireless Ad Hoc Networks” By A.D.Amis, R. Prakash, T. H. P. Vuong and D.T. Huynh, in Proceedings of IEEE INFOCOM, March 2000.

[6] “An Energy efficient pre-schedule scheme for Hybrid CSMA/TDMA MAC in WSN” By Wei Wang, Honggang Wang, IEEE, 2006.

[7] “HYMAC: Hybrid TDMA/FDMA medium access control protocol for WSN” By Mastooreh Salajegheh,, Hamed Soroush, IEEE, 2007.

[8] “Construction Of Decentralized Lifetime Maximizing Tree for Data Aggregation in Wireless Sensor Networks” By Deepali Virmani,Satbir Jain, PWASET VOLUME 40 APRIL 2009 ISSN 2070-3740.

[9] “An Efficient Data Aggregation Scheme Using Degree Of Dependence on Clusters in WSNs” By Tetsushi Fukabori Hidehisa Nakayama, Hiroki Nishiyama, Nirwan Ansari, and Nei Kato Graduate School of Information Sciences,Tohoku University, Sendai, Japn Tohoku Institute of Technology, Sendai, Japan Advanced Networking Lab ECE Department, New Jersey Institute of Technology, Newark, NJ, USA.

[10] “Decentralized Lifetime Maximizing Tree with Clustering for data delivery in Wireless Sensor Networks” By Deepali Virmani, Satbir Jain, Department of CSE, BPIT, GGSIPU, Delhi, Department of CSE, NSIT, DU, delhi.

[11] Deepali Virmani and Satbir Jain “Performance analysis of lifetime maximizing trees for data aggregation in wireless sensor networks” International Journal on Computer Science and Engineering, Vol .3 No.1 Jan 2011.

[12] Wireless Sensor Networks– Technology, Protocols and Application ebook by Kazem Sohraby, Daniel Minoli, Taieb Znati.

[13] “Attribute aware data aggregation using dynamic routing in WSN” By Jiao Zhang, Fengyuan Ren, Tao He, Chuang Lin 2010 IEEE.

[14] “Energy efficient wake up scheduling for data collection and Aggregation” By Yanwei Wu, Member, IEEE, Xiang-Yang Li, Senior Member, IEEE, YunHao Liu, Senior Member, IEEE, and Wei Lou.