

Maximum Lifetime Data Aggregation Problem in Wireless Sensor Networks

KAVINDRA KUMAR
Department of Physics,
Bharat Institute of Technology,
Meerut

Abstract

This paper studies energy efficient routing for maximum lifetime data aggregation problem in wireless sensor networks. We consider a network of energy-constrained sensors that are deployed over a region. Each sensor periodically produces information as it monitors its vicinity. The main process in such a network is the systematic aggregation and transmission of sensed data to a base station for further processing. During data aggregation process, sensors perform in-network aggregation of data packets enroute to the base station. The main challenge in data aggregation is to maximize the relative network lifetime, given the energy constraints of the sensors. Our goal is to maximize the lifetime of a wireless sensor network, given the location of n sensors and a base station together with the available energy at each sensor. We have done simulations for single and multiple base stations to solve the data aggregation problem with aggregation in wireless sensor networks. Our simulation results demonstrate that we obtain better performance than the existing approaches in terms of relative network lifetime.

Keywords: *sensor network, linear programming, aggregation, relative network lifetime*

1. INTRODUCTION

The past few years have seen the rapid proliferation of small wireless devices for personal communication. Such devices are achieving increasing levels of connectivity with large networks such as the Internet, and among themselves as peer-to-peer networks [22]. Alongside these recent advancements in wireless networks, there have been significant developments in low-power digital circuit design, sensing technology, and Micro Electro-Mechanical Systems (MEMS) [23]. The amalgamation of all these technologies has sparked great interest in creating miniature units that combine physical sensing and wireless communication - effectively, a wireless micro-sensor device. Large numbers of such sensor nodes can be disseminated in a region, and can automatically collect and analyze data from the physical environment. A large network of such nodes is collaborating their sensing efforts and it offers significant new opportunities in the study, monitoring, and maintenance of physical environments [8]. Like most wireless systems, sensor networks must effectively manage the power consumption of individual nodes to achieve satisfactory system lifetimes. It is imperative to design low-power hardware and power-aware operating systems so that power savings can be achieved at the node level.

The applications of sensor networks are many. They are used in military, health, environment, underwater, biological, and in many other fields where sensing of the data is required. The order of nodes may vary from a dozen in a home appliance to millions in a military field. However these wireless sensor networks differ from Ad-hoc networks. The differences between sensor networks and ad-hoc networks are outlined below:

- The number of sensor nodes in a sensor network can be several orders of magnitude higher than the nodes in an ad hoc network.
- Sensor nodes are densely deployed.
- Sensor nodes are prone to failures.
- The topology of a sensor network changes very frequently.

- Sensor nodes mainly use broadcast communication paradigm whereas most ad hoc networks are based on point-to-point communications.
- Sensor nodes are limited in power, computational capacities, and memory.
- Sensor nodes may not have global identification (ID) because of the large amount of overhead and large number of sensors.

The protocols which are being currently used for the traditional Ad-hoc Networks can not be used for the sensor networks because:

- The topology of a sensor network changes from time to time.
- Sensor networks use broadcast communication while Ad-hoc networks use point-to-point communication.
- The failure rate of the nodes is quite high.
- Because of large scalability, nodes can not have a global ID as in Ad-hoc networks.
- Power consumption is of critical importance in sensor networks.

As the number of nodes in a sensor networks is very high, multi-hop communication gives more advantage than traditional single hop communication as it consumes less power. Again power consumption of each node is critical, while a compromise is made on Quality of Service [16]. Also the node architecture should be flexible in the sense that the user must have an option of extending the network lifetime at the expense of lower efficiency or higher transmission delay. Recent years have witnessed a growing interest in the application of wireless sensor networks in unattended environments. Nodes in such applications are equipped with limited energy supply and need careful management in order to extend their lifetime. In order to conserve energy, many of the routing protocols proposed for wireless sensor networks reduce the number of transmitted packets by pursuing in-network data aggregation. Almost all of the aggregation schemes presented in the literature strive to save sensor's energy while considering unconstrained data traffic. However, aggregation extends the queuing delay at the relay nodes and can thus complicate the handling of latency-constrained data.

Data fusion or aggregation has emerged as a basic tenet in sensor networks. The key idea is to combine data from different sensors to eliminate redundant transmissions, and provide a rich, multi-dimensional view of the environment being monitored. This paradigm shifts the focus from address-centric approaches finding routes between pairs of end nodes to a more data-centric approach finding routes from multiple sources to a destination that allows in-network consolidation of data [3]. There are several approaches to save the lifetime of the wireless sensor network. Out of those approaches data aggregation turns out to be one of the outstanding approaches which promises to be considerable energy saving, as it has emerged as a basic approach in WSN's in order to reduce the number of transmissions among sensor nodes and base station and hence minimizing the overall power consumption in the network. Data aggregation is affected by several factors such as the placement of aggregation points, the aggregation function, and the density of the sensors in the network. The determination of an optimal selection of aggregation point is extremely important. The number of packets transmitted from one node to another also affects the battery lifetime of the networks. The TinyOS operating system that can be used by an ad-hoc network of sensors locates each other and route data. The implementation of five basic database aggregates, i.e. count, min, max, sum, and average, based on the TinyOS platform and such a generic approach for aggregation leads to significant power (energy) savings. A class of aggregation is particularly well suited to the in-network regime. Such aggregates can be expressed as an aggregate function f over the sets a and b , such that $f(a \cup b) = g(f(a), f(b))$ [18]. We have to minimize the energy expended by the sensors during the process of data gathering. Directed diffusion is based on a network of nodes that can coordinate to perform distributed sensing of an environmental phenomenon. Such an approach achieves significant energy savings when intermediate nodes aggregate responses to queries [9]. The SPIN protocol uses meta-data negotiations between sensors to eliminate redundant data transmissions through the network. In PEGASIS [19], sensors form chains so that each node transmits and receives

directed trees, each rooted at the base station and spanning all the sensors i.e. a schedule has one tree for each round. The lifetime of a schedule equals the lifetime of the system under that schedule.

The solution to this problem is obtained as organized in the following steps given below.

1. We need to find out the edge capacities between different nodes with constraint that the battery lifetime of every node is maximized.
2. Obtain the corresponding aggregation tree for each round.

The edge capacities between different nodes for the given network with the given constraint that the battery lifetime of each node is maximized, is obtained by solving linear programming of that network with that given constraint. After solving the linear program, get the corresponding aggregation trees.

3.2 Linear Program Formulation for WSN

[1] Consider a schedule S with lifetime T rounds. Let $f_{i,j}$ be the total number of packets that node i (a sensor) transmits to node j (another sensor or base station) in S . Since any valid schedule must respect the energy constraints at each other sensor, for $i = 1, 2, 3 \dots, n$, is given by

$$\sum_{j=1}^{n+1} f_{i,j} \cdot T x_{i,j} + \sum f_{j,i} \cdot R x_i \leq E_i \quad (3)$$

Each sensor, for each one of the T rounds, generates one data packet that needs to be collected, possibly aggregated, and eventually transmitted to the base station. The schedule S induces a flow network $G = (V, E)$. The flow network G is a directed graph having as nodes all the sensors and the base station, and having edges (i, j) with capacity $f_{i,j}$ whenever $f_{i,j} > 0$.

Theorem [1]. Let S be a schedule with lifetime T , and let G be the flow network induced by S . Then, for each sensor s , the maximum flow from s to the base station t in $G \geq T$.

The corresponding linear programming for the given constraint is given in [1] in which some other constraints to ensure the flow conservation of the packets at each node are also included.

3.3. Obtaining Aggregation Tree

After solving the linear programming, the aggregation tree for the network is obtained as follows [1]. Given an admissible flow network G with lifetime T and a directed tree A rooted at the base station 't' (A need not span all nodes in G and not necessarily aggregation tree) with lifetime f , we define the (A, f) -reduction G' of G to be the admissible flow network that results from G after reducing by f , the capacities of all of its edges that are also in A , $G' = (A, f)$. This reduction of the G' is possible if maximum flow from a node to the base station 't' in G' is $\geq T - f$ for each vertex 'v' in the G' . Note that A does not have to span all the vertices of G , and thus it is not necessarily an aggregation tree. Moreover, if A is an aggregation tree, with lifetime f , for an admissible flow network G with lifetime T , and the (A, f) -reduction of G is feasible, then the (A, f) -reduced flow network G' of G is an admissible flow network with lifetime $T - f$. Therefore, by using a simple iterative procedure, we can construct a schedule for an admissible flow network G with lifetime T , provided we can find such an aggregation tree A .

GETTREE Algorithm

GETTREE (Flow Network G , Lifetime T , Base Station t)

- initialize $f \leftarrow 1$
- let $A = (V_0, E_0)$ where $V_0 = \{t\}$ and $E_0 = \phi$
- while A does not span all the nodes of G do
- for each edge $e = (i, j) \in G$ such that $i \notin V_0$ and $j \in V_0$ do
- let A' be A together with the edge e

- // if the $(A',1)$ -reduction of G is feasible
- let G_r be the $(A',1)$ - reduction of G
- if $\text{MAXFLOW}(v,t,G_r) \geq T-1$ for all nodes v of G
- // replace A with A'
- $V_0 \leftarrow V_0 \cup \{i\}, E_0 \leftarrow E_0 \cup \{e\}$
- break
- let c_{\min} be the minimum capacity of the edges in A
- let G_r be the (A, c_{\min}) -reduction of G
- if $\text{MAXFLOW}(v,t,G_r) \geq T - c_{\min}$ for all nodes v of G
- $f \leftarrow c_{\min}$
- replace G with the (A, f) -reduction of G
- return f, G, A

The GETTREE algorithm is used to get an aggregation tree A with lifetime $f \leq T$ from an admissible flow network G with lifetime T . Throughout this algorithm, we maintain the invariant that A is a tree rooted at t and the (A, f) -reduction of G is feasible. Tree A is formed as follows. Initially A contains just the base station. While A does not span all the sensors, we find and add to A an edge $e = (i, j)$, where $i \notin A$ and $j \in A$, provided that the (A', f) -reduction of G is feasible. Here A' is the tree A together with the edge e , and f is 1 or the minimum of the capacities of the edges in A' .

3.4. Obtaining the Schedule

Next, how to get a schedule from an admissible flow network is discussed. A schedule is a collection of directed trees that span all the sensors and the base station, with one such tree for each round. Each such tree specifies how data packets are gathered and transmitted to the base station. These trees are called aggregation trees. An aggregation tree may be used for one or more rounds; the number of rounds is indicated by f , that an aggregation tree is used, by associating the value f with each one of its edges; f is called to be the lifetime of the aggregation tree.

Finally, a collection of aggregation trees from an admissible flow network G with lifetime T can be computed by using the GETSCHEDULE algorithm as given below, such that T data packets from each of the sensors are aggregated and transmitted to the base station t .

GETSCHEDULE Algorithm

GETSCHEDULE (Flow Network G , Lifetime T , Base Station t)

```

S ← 0
while T > 0 do
    [f, G, A] ← GETTREE(G, T, t)
    S ← S ∪ {A}
    T ← T - f
return S
```

This means that the GETSCHEDULE algorithm can always find a sequence of aggregation trees that can be used to aggregate and transmit T data packets from each sensor to the base station.

21. Madden, S., Szewczyk, R., Franklin, M. J., and D. Culler, Supporting Aggregate Queries Over Ad-Hoc Wireless Sensor Networks, Proceedings of 4th IEEE Workshop on Mobile Computing and Systems Applications, 2002.
22. Min, R., Bhardwaj, M., Cho, S.H., Sinha, A., Shih, E., Wang, E., and Chandrakasan, A.P., Low-Power Wireless Sensor Networks, VLSI Design, 2001.
23. Rabaey, J., Ammer, J., da Silva Jr, J.L. and Patel, D., Pico Radio: Ad-hoc Wireless Networking of Ubiquitous Low-Energy Sensor/Monitor Nodes, Proceedings of the IEEE Computer Society Annual Workshop on VLSI, 2000.
24. Singh, S., Woo, M. and Raghavendra, C., Power-aware Routing in Mobile Ad Hoc Networks, Proceedings of 4th ACM/IEEE Mobicom Conference, 1998.
25. Chvatal Vasek, Linear Programming, W. H. Freeman and Company, 1983.
26. Bixby Robert E., Implementing the Simplex Method: The Initial Basis, ORSA Journal on Computing, Vol. 4, No. 3, 1992.
27. Andersen Erling D. and Knud D. Andersen, Presolving in Linear Programming, Mathematical Programming, Vol. 71, pp. 221-245, 1995.
28. Kim, D., Park, J., Toh, C.K and Yanghee Choi, Power aware Route Maintenance Protocol for Mobile Ad hoc Networks, IEEE Personal Communications 2003, Vol. 22, pp 501-506.
29. Rodoplu, V. and Meng, T., Minimum Energy Mobile Wireless Networks, IEEE, JSAC, vol. 17, no. 8, Aug. 1999, pp. 1333-44.
30. Stojmenovic, I., and Lin, X., Power-aware localized routing in wireless networks, Proceedings IEEE IPDPS, Cancun, Mexico, May 2000, pp. 1123-1133.
31. Lindsey, S., Raghavendra, C. S., and Sivalingam, K., Data Gathering in Sensor Networks using the Energy*Delay Metric, Proceedings of the IPDPS Workshop on Issues in Wireless Networks and Mobile Computing, 2001

7. FIGURES

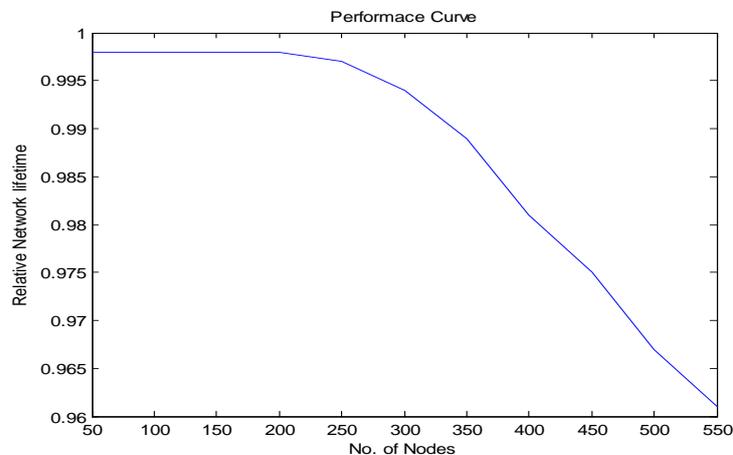


Figure 1. Relative Network Lifetime Vs No. of Nodes One BS)

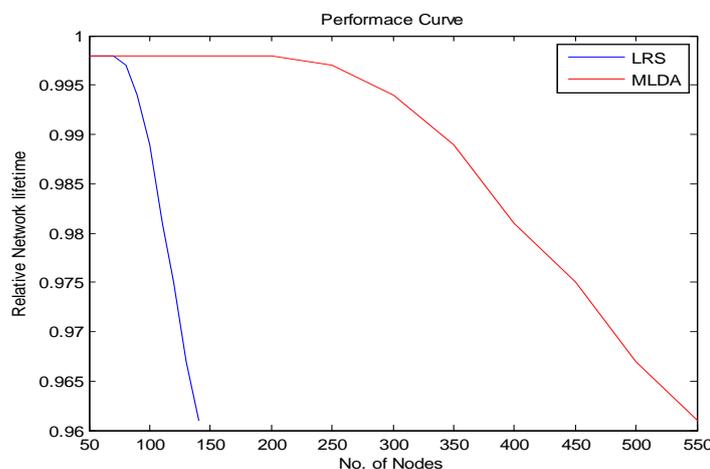


Figure 2. Relative Network Lifetime Vs No. of Nodes (MLDA One BS vs. LRS)

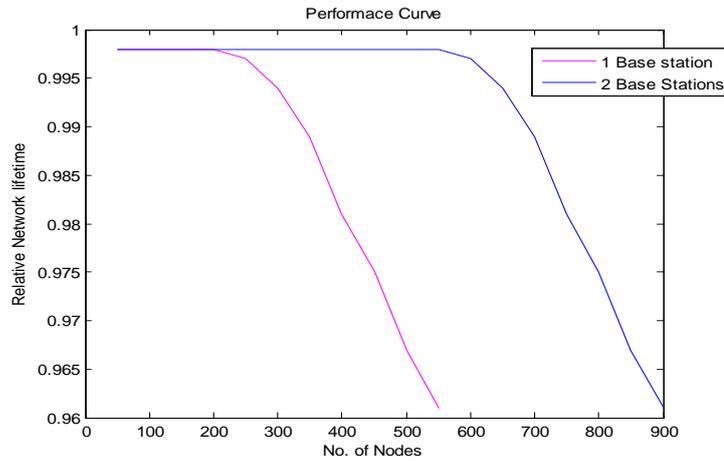


Figure 3. Relative Network Lifetime Vs No. of Nodes (Two BS)

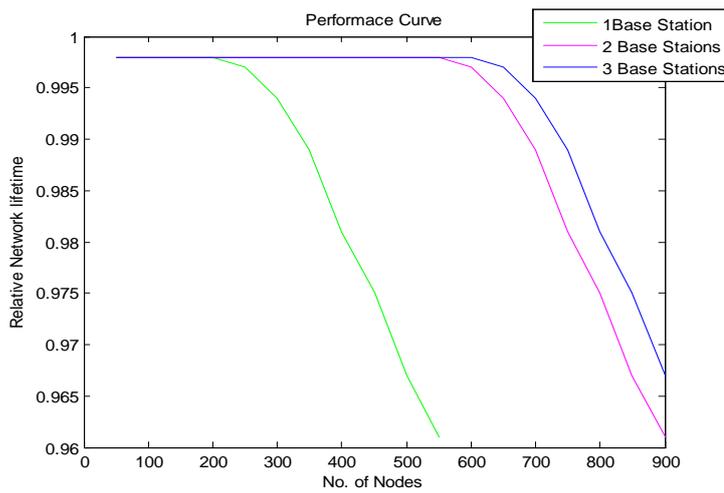


Figure 4. Relative Network Lifetime Vs No. of Nodes (Three BS)

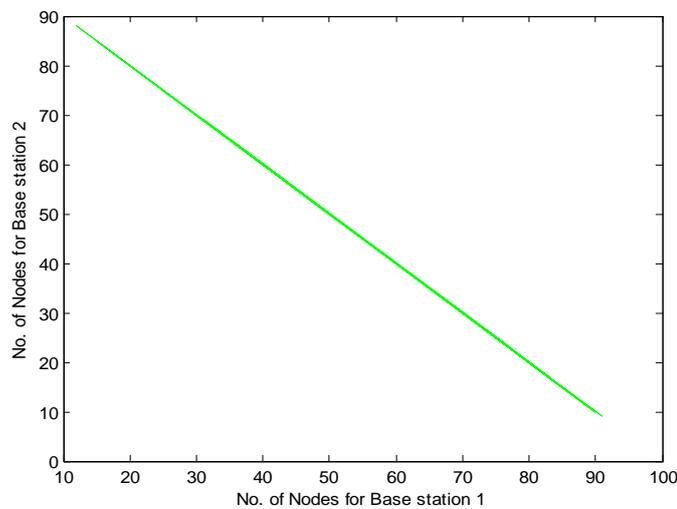


Figure 5. Base station placement (Two BS)

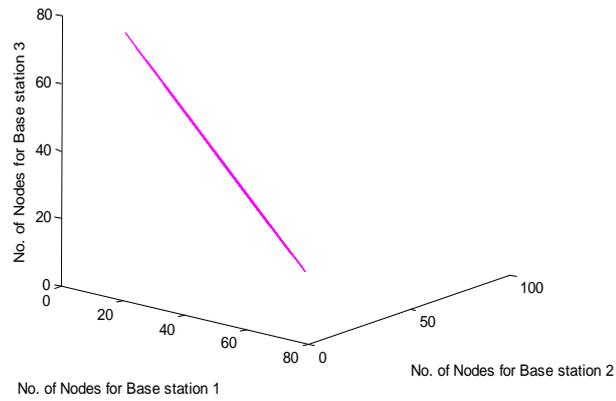


Figure 6 (a). Plot for the three base stations for placement in one angle

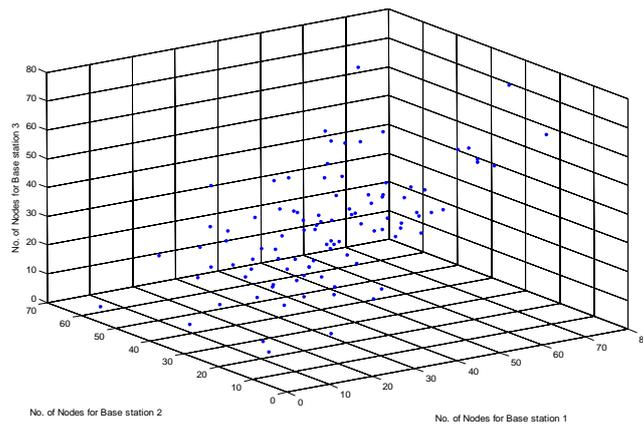


Figure 6(b). Plot for three base stations placement in another angle