Data Aggregation in Wireless Sensor Networks

Arabind Kumar, Ram Baksh, Ranjan Kumar Thakur, Aditya Pratap Singh

Delhi University, Department of Mathematics, Kamala Nehru College, August Kranti Marg, New Delhi - 110049

Delhi University, Departmento of Mathematics, Dyal Singh College, Lodhi Road, New Delhi - 110003

Delhi University, Department of Mathematics, Deshbandhu College, Kalkaji, New Delhi - 110019

Delhi University, Department Of Mathematics, PGDAV College(Evening), Nehru Nagar, New Delhi - 110065

Abstract: *Wireless sensor networks consist of sensor nodes with sensing and communication capabilities. Efficient aggregation of data collected by sensors is crucial for a successful application of wireless sensor networks (WSNs). Both minimizing the energy cost and reducing the time duration of data aggregation have been extensively studied for WSNs. Algorithms with theoretical performance guarantees are only known under the protocol interference model, or graph-based interference models generally. A fundamental challenge in the design of Wireless Sensor Network (WSNs) is to maximize their lifetimes. Data aggregation has emerged as a basic approach in WSNs in order to reduce the number of transmissions of sensor nodes, and hence minimizing the overall power consumption in the network.*

Keywords: Dominating, Aggression, Protocol, WSN

1. Introduction

Wireless Sensor Networks are a result of the combination of advances made in the field of analog and digital circuitry, wireless communications and sensor technology. A wireless sensor network typically consists of small devices called sensor nodes that are capable of sensing the environment around them. The sensor nodes are devices that are capable of sensing, gathering, storing and transmitting information. The main advantage of these nodes is their self-organizing capability. Large networks of such small nodes are therefore growing in use. The sensor nodes can be deployed anywhere without actually having to install or deploy them manually. In remotely inaccessible areas, these sensor nodes are just strewn across the desired sensor field. The self-organizing capability of the sensor nodes enables the nodes to form a cooperative network and gather information. This information can then be retrieved. Thus, sensor networks enable intelligent monitoring of inaccessible areas with ease and accuracy. Wireless sensor network (WSN) is a term used to describe an emerging class of embedded communication products that provide redundant, fault-tolerant wireless connections between sensors, actuators and controllers [1].

The main challenges with wireless sensor network are how to provide maximum lifetime to network and how to provide robustness to network. As sensor network totally rely on battery power, the main aim for maximizing lifetime of network is to conserve battery power or energy. In sensor network, the energy is mainly consumed for three purposes: data transmission, signal processing, and hardware operation. The 70% of energy consumption is due to data transmission. So for maximizing the network lifetime, the process of data transmission should be optimized. The data transmission can be optimized by using efficient routing protocols and effective ways of data aggregation.

2. Tree-Based Data Aggregation

In this all nodes are organized in form of tree means hierarchical, with the help of intermediate node we can perform data aggregation process and data transmit leaf node root node. Tree-based data aggregation is suitable for applications which involve in-network data aggregation. An example application is radiation-level monitoring in a nuclear plant where the maximum value provides the most useful information for the safety of the plant. One of the main aspects of tree-based networks is the construction of an energy efficient data-aggregation tree.

2.1 Data Aggregation by Tree Construction

Consider a given network G , V is a set of all sensor nodes that consist the network G and r is distance between two nodes (u, v) that is length between two nodes .The combination form by these link is denoted by G(V, r) .

3. Definitions

3.1 Center of Topology
Any node v_0 is called the center of topology of G(V, r) if v_0 =min dg(u, v) where dg(u, v) is the hope distance between nodes u and v in G(V, r).

3.2 Connecting Dominating Set (CDS)

Consider any graph G= (V, E), a subset V_0 of V is a dominating set if each node in V is either in V_0 or adjacent to some node in V_0 . Nodes in V_0 are called dominators where as nodes not in V_0 are called dominatees. A subset C of V is called a connected dominating set if C is a dominating set and C includes a connected sub graph.

3.3 Maximal Independent Set (MIS)

The collections of nodes that are separated at two hop distance [10]. Construct the aggregation tree on the

communication graph G (V, r) using Algorithm1. The basic idea of Algorithm 1 is to construct a tree similar to the breadth-first-search tree, with the following properties:

(1) The depth of the tree is within a small constant factor of the diameter D (G),

(2) Each internal node will be connected with at most a constant number of other internal nodes.

The second property ensures that we can schedule the transmissions of internal nodes in constant time-slots. The algorithm based on the following properties.

(1) We select the topology center of G (V, r) as the root of our BFS tree. The topology center selection enables us to reduce the latency to a function of the network radius R(G),instead of the network diameter D(G)

(2) After the topology center gathered the aggregated data from all nodes, it will then send the aggregated result to the sink node via the shortest path from the topology center to v_0 the sink node.

In a tree based network, sensor nodes are organized into a tree where data aggregation is performed at intermediate nodes along the tree and a concise representation of the data is transmitted to the root node. Tree based data aggregation is suitable for applications which involve in-network data aggregation. An example application is radiation level monitoring in a nuclear plant where the maximum value provides the most useful information for the safety of the plant. One of the main aspects of tree-based networks is the construction of an energy efficient data aggregation tree.

4. A New Approach for Finding Maximal Independent Set (MIS)

According to the definition of maximal independent set, it is the collection of the nodes which are separated at two hop distance from a certain node. Let G be a non cyclic graph and V be the set of all nodes. In this approach for finding MIS, I am using two function find_MIS() and get_node() and call these function in the main algorithm.

In function find_MIS() first of all pass a center node and then storing all the adjacent node of center in the set S and then for each node of the set S find rank and visit of node and if the rank of node is 2 and node is not center and visit of node is less than or equal to two then print the node. Repeat the same process for each node of graph. The function get_node() is worked on the set S. This function is return the first node of set S and the decrease the one in the number of nodes in the set S.

Algorithm:

```
begin
    n ← no. of nodes in G
    V ← set of nodes = {0,1,...,n−1}
    for i ← 0 to n−1
        rank (i) ← 0
        visit (i) ← 0
    end for
    for i ← 0 to n−1
        find _ MIS (i)
    end for
end
```

```
find _MIS(node)
begin
    centre ← node
    S ← set of all nodes adjacent to centre
    C ← centre of graph and Display(C)
    while (S is non−empty)
        do
        { node ← get _ Node(S)
            rank(node) ← rank(centre)+1
            visit(node) ← visit(node)+1
            if (rank(node) = 2 & &node ≠ C & &visit(j) ≤ 2)
                Display(node)
                rank(node) ← 0
            endif
        }
    end while
```

```
get _ Node(S)
begin
    n(S) ← no. of nodes in S
    return ( first node of S)
    n(S) ← n(S)−1
end
```

5. Simulation Result for the Algorithm Aggregation Tree Construction

The n nodes randomly deploy in a two dimensional square area and randomly choose one node as the sink node. The objective is to measure the latency for the sink node to get the aggregation result of all data using our aggregation scheduling algorithm. Note that latency is defined as the number of time slots needed to aggregate all data from within the network to the sink node. The effect of network radius R, we create a network instance similar to the first one with two crucial modifications we vary the network deployment area from (200m x200m) to (1000m x1000m) gradually with the network size . In other words, fix the node density in the network deployment area. It is find that Δ is nearly fixed (around25) as well in the corresponding communication graph. At the same time, R increases monotonously with the network size. Now measure the performance of Algorithm under this condition; the average performance is illustrated in Figure. This is because we use Δ+R as a lower bound on the latency. This means that either Algorithm can be potentially further improved or the lower bound Δ+R should be improved for data aggregation under the physical interference model.

6. Conclusion of Algorithms

In first algorithm each child node send their data to parent node and finally all the data are collected at the center node and then center node send data to sink or base station. Since

in this algorithm for traversing of graph, using breath first search, the benefit is that the data are send from different node to center node by shortest path by this way the latency decreases and node save some part of energy and finally life time of sensor networks are increased.

References

[1] Andrew S. Tanenbaum, "Computer Networks", Fourth Edition, *Pearson Prentice Hall*, 2003.

[2] Bo Yu, Jianzhong Li and Yingshu Li, "Distributed Data Aggregation Scheduling in Wireless Sensor Networks", *IEEE INFOCOM*, pp. 2159 – 2167, 2009.

[3] Carlos de Morais Cordeiro and Dharma Prakash Agrawal, "Ad Hoc & Sensor Networks", *Cambridge University Press*, 2010.

[4] Hani Alzaid, Ernest Foo and Juan Gonzalez Nieto, "Secure Data Aggregation in Wireless Sensor Network: A survey", *Australasian Information Security Conference*, Wollongong, Australia, vol. 81, pp. 93 – 105, 2008.

[5] Jamal N. Al-Karaki, Raza Ul-Mustafa and Ahmed E. Kamal, "Data Aggregation in Wireless Sensor Networks – Exact and Approximate Algorithms", *High Performance Switching and Routing*, pp. 241 – 245, 2004.

[6] Jie Gao, Leonidas Guibas and John Hershberger, "Sparse Data Aggregation in Sensor Networks", *Information processing in sensor networks*, pp. 430-439, 2007.

[7] Jia Guo, Jian'an Fang and Xuemin Chen, "Survey on Secure Data Aggregation for Wireless Sensor Networks ", *Service Operations, Logistics and Informatics*, pp. 138 – 143, 2011.

[8] Mukesh Kumar Jha,T.P. Sharma, "Secure Data aggregation in Wireless Sensor Network: A Survey", *International Journal of Engineering Science and Technology*, vol. 3, pp. 2013 – 2019, 2011.

[9] Nandini S. Patil, Prof. P. R. Patil, "Data Aggregation in Wireless Sensor Networks", *International Conference on Computational Intelligence and Computing Research*, 2010.

[10] P.-J. Wan, K. M. Alzoubi and O. Frieder, "Distributed Construction of Connected Dominating Set in Wireless ad hoc Networks", *IEEE INFOCOM,* vol.3, pp. 1597 – 1604, 2002.