A Graph-Center-Based Scheme for Energy-Efficient Data Collection in Wireless Sensor Networks

Dajin Wang

Department of Computer Science
Montclair State University,
Upper Montclair, NJ 07043, USA
wang@pegasus.montclair.edu

Abstract. We consider the problem of sensor data collection in a wireless sensor network (WSN). The geographic deployment of sensors is random, with an irregular network topology. We propose a data collection scheme for the WSN, based on the concept of the *center of the graph* in graph theory. The purpose of the scheme is to use less power in the process of data collection. Because it is mostly true that the sensors of WSN are powered by batteries, power saving is an especially important issue in WSN. In this paper, we will propose the energy-saving scheme, and provide the experimental results. It is shown that under the energy consumption model used in the paper, the proposed scheme saves about 20% of the power collecting data from sensors.

Keywords: Energy efficiency, Graph center, Hierarchical structures, Wireless networks, Wireless Sensor Networks.

1 Introduction

A Wireless Sensor Network (WSN) is composed of a large number of sensor nodes, and one (or a few) "central" node(s). The sensors are deployed in various physical environments mainly for the collection of physical world data. The data are transmitted to, or gathered by, the central nodes for aggregation, analysis, and processing. The central nodes also play the role of manager of the WSN. The communication among nodes is all via wireless means. Therefore all nodes are equipped with radio transceivers/receivers. WSNs have very promising prospect in many applications, such as environment monitoring, traffic monitoring, target tracking, and fire detection.

Different models of WSN have been proposed. However some basic characteristics can be observed that are common in most proposed models.

- They are all composed of a large number of sensor nodes, and a small number of master nodes (a.k.a. central nodes, or base station);
- All sensor nodes are relatively low cost, perform relatively limited computational operation. Their main job in the whole system is to collect raw

J. Cao et al. (Eds.): MSN 2006, LNCS 4325, pp. 579–587, 2006.

data, and render it to the master nodes, with or without some primitive preprocessing;
- The master node(s) collect the data from all sensors, and analyze/process them. They are more powerful, costlier processors than ordinary sensors. The master nodes are also the managers of the network.

A WSN can have either just one central node or a group of central nodes, depending on the network's scale of geographical coverage and/or cost effectiveness consideration. In a single-center WSN, the central node, a.k.a. *base station*, collects and processes data from all sensors. It is also the sole manager of the entire network system. In a multicenter WSN, the tasks of data collection, aggregation, processing, and network management are distributed among a group of nodes working collaboratively. The organization of these master nodes is one of the essential issues in the design of WSN architecture.

There are many different WSN models. Topologically speaking, a WSN can be of *regular* or *arbitrary* topology. One example of regular topological structure is the COSMOS model (Cluster-based heterOgeneouS MOdel for Sensor networks) proposed in [8]. COSMOS is a cluster-based, hierarchical model for WSN. It comprises of a large number of low power, low cost sensors, presumably distributed in a large physical environment. The distribution of sensors is close to uniform. That is, in each unit area there is a sensor with high likelihood. Sensors are organized into equal-sized, square-shaped clusters according to their spatial proximity. For each sensor cluster, there is a clusterhead. Sensors within a cluster communicate in a time synchronized manner, using single hop communication. The clusterheads form a mesh-like topology and communicate asynchronously. In an WSN of arbitrary topology, sensors are deployed in a random manner. The network can then be modeled by a graph $G = (V, E)$. Each node in V represents a sensor. Each edge in E linking nodes u and v represents communication between u and v via wireless means. One of the nodes is designated as the base station of the WSN.

In this paper, we consider the problem of collecting data, from all sensor nodes to the base station, in an energy-saving manner. Using the COSMOS model's hierarchical idea, our proposed scheme applies a hierarchical, two-phase approach to the arbitrary topology. That is, the WSN is divided into logical hierarchies. The lower level sensors are grouped into *clusters*, and a *clusterhead* collects data within the cluster. The collected data are aggregated, preprocessed, and then forwarded to the base station. The purpose of the scheme is to use less power in the process of data collection. By the nature of WSN, all sensors are supposed to be powered by batteries. Therefore energy preservation is an especially crucial issue in WSN. The experiments show that under the energy consumption model used in the paper, the proposed scheme can save about 20% of the power collecting sensor data.

The rest of this paper is organized as follows. In Section 2, we describe the WSN model we will be working on. In Section 3, we will present the clustering scheme minimizing energy consumption. The scheme is based on the concept of the center of the graph in graph theory. Section 4 presents simulation results to demonstrate proposed scheme's gain in energy saving. Section 4 also gives concluding remarks and discusses possible directions the work of this paper can be extended.

2 The Sensor Network Model

A wireless sensor network resembles a conventional parallel and distributed systems in many ways. However, several unique characteristics standout to call for redefinition, or modification, of the network model. Those characteristics include energy efficiency consideration, communication reliability, and global awareness of individual nodes, among others. Because of the wide diversity of sensor applications, it is hard to capture all characteristics in one single model.

In this paper, we consider the WSN with its sensors randomly deployed, without following any proximal patterns. A sensor communicates with another one via radio transmitter/receiver. If a node needs to transmit to another node out of its radio range, the message has to be relayed by intermediate nodes. Such a WSN can be readily modeled with a graph $G = (V, E)$, illustrated in Figure 1. Each node in V represents a sensor. There is a link $(u, v) \in E$ if and only if sensors u, v are in each other's radio transmission range (Figure 1 (a)). In Figure 1, the primed letters (a', c' etc.) on the radio circle identify the sensors they belong to. Figure 1 (b) is an example WSN of seven sensors and its corresponding graph.

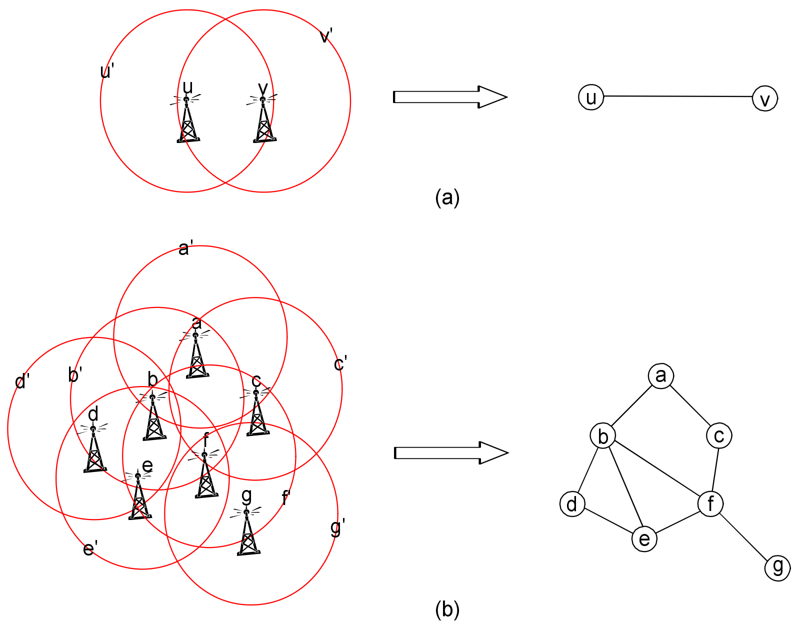

Fig. 1. (a) Node/edge definition in the graph model; (b) An example set of randomly deployed sensors, and its corresponding graph

For the purpose of power preservation, we assume that all sensors use as little power as possible for radio transmission, so that the transmission range covers just a few neighboring sensors. We also assume that there is only one transmission range, as opposed to multi-range models in some literatures. If a sensor wants to send message/data to the base node, it can only do so by relaying through intermediate sensors (routing scheme in this context is another issue, which will not be addressed in this work). We also assume a connected graph. That is, we do not consider isolated sensors or components in the WSN.

We quantify the energy dissipated by one hop of sensor transmission to a normalized unit. Refer to the example in Figure 1 again: If sensor a wants to send one piece of data to sensor g, at least 3 hops are needed; therefore 3 units of energy will be consumed, e.g. 1 unit for transmission from a to c, 1 from c to f, and 1 from f to g. In the discussion of the following section, we only consider data relaying via a shortest path.

3 A Hierarchical Scheme for Energy-Efficient Data Collection

3.1 The Designation of Base Station

We'd like to designate such a sensor as the base station, that it uses the least amount of total power to collect data from all sensors. To formulate the problems quantitatively, we first assume a model for calculating power consumption. It should be pointed out that the model is a simplified abstraction from vastly variable real scenarios. Refer to Figure 1 (b) again. We use the number of relaying hops to represent needed power to transmit data from sensor to base. The farther the sensor, the more hops are needed to relay the data, and the more power is consumed. Secondly, to measure the saving in power, we focus on the scenario of base station collecting one unit of data from each sensor. A natural choice would be to pick a "central node" of the underlying graph G. The central node(s) of an arbitrary graph can be established through the following definitions.

Definition 1. *In an undirected graph $G = (V, E)$, the eccentricity of a node v is the greatest distance between v and any other node.*

Definition 2. *The radius of a graph G is the minimum eccentricity of any node in G.*

Definition 3. *The center of a graph G is the set of nodes of G whose eccentricity is equal to the radius.*

In the examples of Figure 2, the number by a node is the node's eccentricity, and the grey nodes constitute the center of a graph. It can be observed that the notion of "center nodes" is based on the idea that these nodes have the shortest distance to all other nodes. So it would be appropriate to designate one of the center nodes to be the base station. In the examples of Figure 2, the circled nodes are designated as base station. We use the total number of hops

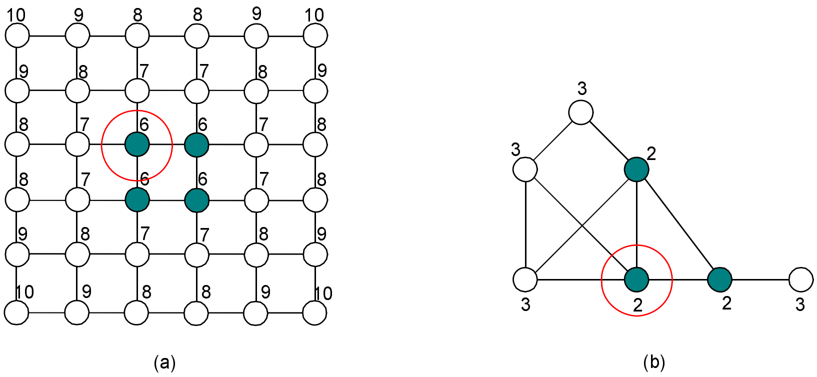

Fig. 2. Two example graphs: Eccentricities and centers

to represent power incurred in the process of data collection. To calculate power consumption for collecting one unit of data from all sensors, we just add up the distances from all sensor nodes to the base node. In Figure 2, the total power consumption for the two example WSNs are 108 in (a), and 8 in (b), respectively.

Having just one node performing the function of base for the entire WSN would be ideal. However it might not be feasible as the size of WSN grows larger, and the geographic range wider. Issues such as energy limitation, energy balancing, and scalability make a single-base WSN not only unfavorable, but also difficult to implement. The proposal of hierarchical organization of WSN [8,10] is to distribute computational and managerial tasks to a group of *clusterheads*. The approach will reduce the communication traffic in network, as well as the overall power consumption.

3.2 The Hierarchical Clustering Scheme

In the hierarchical approach, the whole WSN is divided into a set of smaller network clusters. There is still a base station for the whole WSN, chosen from the center nodes as defined in Definition 3. The data collection of base station is now performed in two phases. In the first phase, all clusterheads collect data from sensors in their own clusters. The data is aggregated and/or preliminarily processed in clusterheads. In the second phase, the WSN's base station collects data from all clusterheads. Figure 3 illustrates the structure of hierarchical WSN.

In Figure 3, the whole network is divided into $|C|$ subnetworks (clusters), where C is the set of center nodes as defined in Definition 3. *Each subnetwork of c_i consists of nodes that are closer to c_i than to any other center nodes.* Each subnetwork will then have a center node according to Definition 3, which will act as the clusterhead of the respective cluster. The $|C|$ clusterheads form a network at the upper hierarchy. At the center of upper hierarchy is the WSN's base station. Two examples are illustrated in Figure 4. In Figure 4 (a) and (b), the original networks have 4 and 3 center nodes, respectively. So they are divided

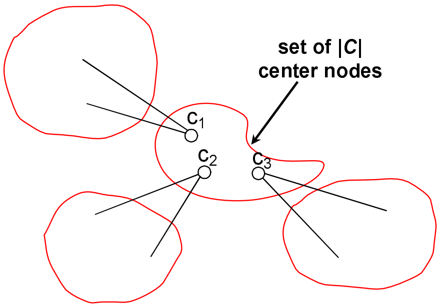

Fig. 3. Hierarchical division of WSN

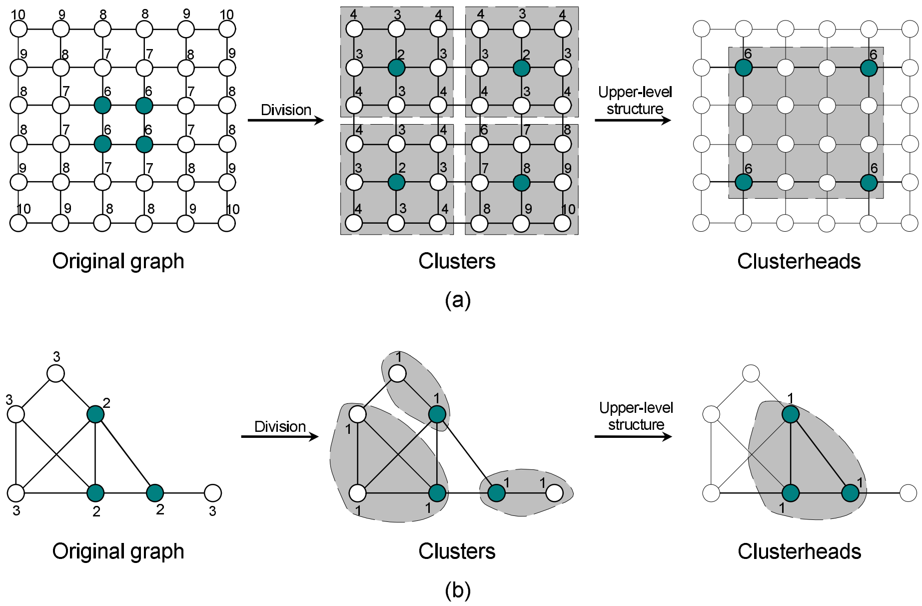

Fig. 4. Examples of hierarchical division

into 4 and 3 clusters. The numbers by the clustered nodes represent their new eccentricities *within* the cluster. The far-right of Figure 4 shows the upper-level of the hierarchy, composed of the clusterheads. The number by a clusterhead is its eccentricity in the upper-level graph. In both examples of Figure 4, the upper-level eccentricities all happen to be the same, but this is not true in general. An upper-level central node can be determined by those upper-level eccentricities, which will be designated as the base node for the whole WSN.

To calculate the total hops incurred in collecting one round of data from all sensors, first in all subgraphs, get the sum of hops from sensors to the corresponding clusterheads. Denote the total sums in all subgraphs as $Cost_I$. Then in the second phase, get the sum of hops from the upper-level nodes to the center node. Denote the sum as $Cost_{II}$. The total cost for one round of data collection is $Cost_I + Cost_{II}$. For the examples in Figure 4 (a) and (b), the total costs of the two-level approach are 60, and 6, respectively. Comparing with the single-level approach, the power-saving rates are 44% and 25%, respectively.

We summarize the steps of hierarchical scheme in the description below.

Cost computation for two-level approach

1. Calculate *eccentricities* for all nodes.
2. Find out all nodes with minimum eccentricities. The subgraph induced by these nodes is called the *center* of G, denoted $C(G)$.
3. Divide G into $|C(G)|$ subgraphs – a node v belongs to a center node $c' \in C(G)$ if $d(v, c') = \min\{d(v, c)|c \in C(G)\}$. If there are more than one such c', then pick any one.
4. After division, calculate eccentricities and centers in all subgraphs.
5. In each subgraph, calculate the cost of every node to a local center node (the sum of hops along the path), and then get sum of all costs. This is the cost for the subgraph.
6. Get the sum of costs of all subgraphs. Call it $Cost_I$ (level-one cost).
7. Calculate the center of C_I, where C_I is the set of all (level-one) center-nodes in subgraphs. Call the level-two center c_{II}. Calculate $\sum_{v \in C_I} d(c_{II}, v)$. Call it $Cost_{II}$ (level-two cost).
8. Total cost: $Cost_I + Cost_{II}$.

---

4 Simulation Result and Concluding Remarks

For irregularly connected networks, simulation is an effective instrument to quantify the competence of a proposed hierarchy method. In the preceding section, we have proposed an energy-efficient clustering scheme for Wireless Sensor Networks, based on the center nodes of a network's underlying graph. Figure 5 shows the ratio of costs for the proposed hierarchical approach vs. the single base-station approach. The scenario being simulated is one round of data collection from all sensor nodes of the network. The cost is the total number of hops incurred in the process.

For the simulation, randomly connected graphs of different sizes are generated, and the corresponding costs for the two approaches are computed and compared. Graphs of 10 nodes through 70 nodes are simulated. The cost for a specific size is the average cost for many random graphs of that size. We can see from Figure 5 that the average cost for the hierarchical scheme is around 80% that of the single base-station approach. In other words, about 20% hops can be saved using the proposed hierarchical scheme for data collection.

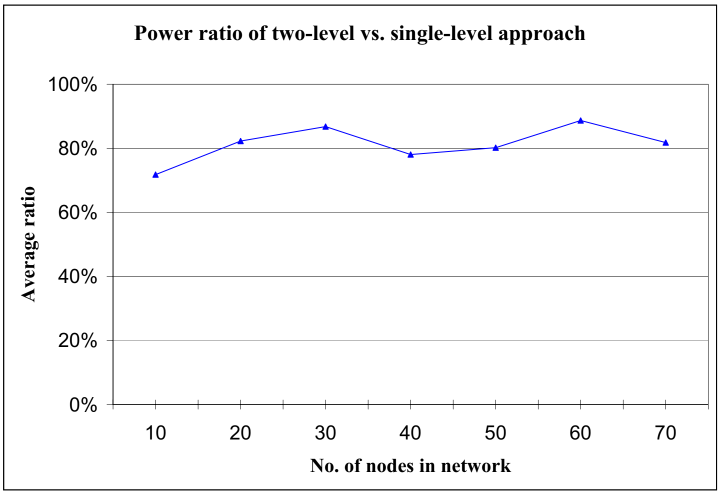

Fig. 5. Simulation result

It is worth pointing out that the proposed scheme is just a "soft" protocol for the task of data collection. It does not impose any hierarchical structure on the original WSN. For other applications, different protocols can be employed on the same network.

Power conservation is a problem that has been extensively addressed in the research of wireless networks, where many open problems exist regarding this issue. We can see some obvious directions to which the work of this paper can be immediately extended. For example, the proposed scheme only considers the saving of *total* transmission hop counts for the *entire* WSN. It does not address the issue of power balancing among sensors. That is, the closer a sensor to a clusterhead, the more power it consumes, because it relays more data packets to the clusterhead. Another worthwhile topic for further research is that in this paper, we assumed a rather simple communication model, not only for tractability reason, but also for the lack of a statistical model that better reflects the realistic transmission activities. Finding an appropriate communication model that's more realistic as well as facilitating tractability would greatly increase the practical relevance of the hierarchical scheme.

References

1. J. Cao and F. Zhang, "Optimal configuration in hierarchical network routing," *Proc. 1999 IEEE Canadian Conference on Electrical and Computer Engineering*, Edmonton, Alberta, Canada, May 1999, pp. 249-254.
2. T. Chu and I. Nikolaidis, "Energy efficient broadcast in mobile ad hoc networks," *Proceedings of Ad-Hoc Networks and Wireless*, 2002, pp. 177-190.

3. R. Marks II, A. Das, M. El-Sharkawi, P. Arabshahi, and A. Gray, "Minimum power broadcast trees for wireless networks: optimizing using the viability lemma," *Proceedings of IEEE International Symposium on Circuits and Systems*, 2002, pp. 245-248.
4. O. Egecioglu and T. Gonzales, "Minimum-energy broadcast in simple graphs with limited node power," *Proceedings of lASTED International Conference on Parallel and Distributed Computing and Systems*, 2001, pp. 334-338.
5. X.-Y. Li and P.-J. Wan, "Constructing minumum energy mobile wireless networks," *ACM Journal of Mobile Computing and Communication Reviews*, Vol. 5, No. 4, 2001, pp. 55-67.
6. S. Lindsey and C. Raghavendra, "Energy efficient broadcasting for situation awareness in ad hoc networks," *Proceedings of International Conference on Parallel Processing*, 2001, pp. 149-155.
7. E. Lloyd, R. Liu, M. Marathe, R. Ramanathan, and S. Ravi, "Algorithmic aspects of topology control problems for ad hoc networks," *Proceedings of Annual Workshop on Mobile and Ad Hoc Networking and Computing*, 2002, pp. 123-134
8. Mitali Singh and Viktor K. Prasanna, "A hierarchical model for distributed collaborative computation in wireless sensor networks," *International Journal of Foundations of Computer Science*, 2004, Vol. 15, No. 3, pp. 485-506.
9. J. Wieselthier, G. Nguyen, and A. Ephremides, "On the construction of energy-efficient broadcast and multicast trees in wireless networks," *Proceedings of IEEE Infocom*, 2000, pp. 585-594.
10. M. Younis, M. Youssef, and K. Arisha, "Energy-aware routing in cluster-based sensor networks," *Tenth ACM Int'l Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems*, Fort Worth, TX, Oct. 2002.