Impact of Network Density on Data Aggregation in Wireless Sensor Networks

Chalermek Intanagonwiwat
USC/ISI
intanago@isi.edu

Deborah Estrin
UCLA and USC/ISI
destrin@cs.ucla.edu

Ramesh Govindan
ICIR and USC/ISI
govindan@isi.edu

John Heidemann
USC/ISI
johnh@isi.edu

Abstract

*In-network data aggregation is essential for wireless sensor networks where energy resources are limited. In a previously proposed data dissemination scheme (*directed diffusion with opportunistic aggregation*), data is opportunistically aggregated at intermediate nodes on a low-latency tree. In this paper, we explore and evaluate* greedy aggregation*, a novel approach that adjusts aggregation points to increase the amount of path sharing, reducing energy consumption. Our preliminary results suggest that, under investigated scenarios, greedy aggregation can achieve up to 45% energy savings over opportunistic aggregation in high-density networks without adversely impacting latency or robustness.*

1 Introduction

Large-scale networks of small and inexpensive sensors may require novel data dissemination paradigms which are scalable, robust, and energy-efficient [2]. One such paradigm is directed diffusion [5] which incorporates data-centric routing and application-specific processing inside the network (*e.g.,* data aggregation). Given that the communication cost is several orders of magnitude higher than the computation cost [7], directed diffusion can achieve significant energy savings with in-network data aggregation [5]. This benefit of data aggregation has been confirmed theoretically [6] and experimentally [3].

The instantiation of directed diffusion described in the earlier work establishes low-latency paths between sources (sensor nodes that detect phenomena) and sinks (user nodes) using only localized algorithms. Paths from different sources to a sink form an *aggregation tree* rooted at the sink. Data from different sources is *opportunistically* aggregated. Whenever similar data happens to meet at a branching node in the tree, the copies of similar data are replaced by a single message. Energy-wise, opportunistic aggregation on a low-latency tree is not optimal because data may not be aggregated (or reduced) near the sources. In this paper, we propose using a greedy incremental tree (GIT) to improve path sharing for more energy savings. We have implemented this *greedy-tree* approach in ns-2 and we compare it to our prior *opportunistic* approach (Section 3). Due to space constraints, some of the more detailed algorithms, simulations, and analysis have been omitted. Please refer to [4] for more details.

2 Greedy Aggregation

Greedy aggregation differs from opportunistic aggregation in path establishment and maintenance. To construct a greedy incremental tree, a shortest path is established for only the first source to the sink whereas each of the other sources is incrementally connected at the closest point on the existing tree.

Path Establishment: Directed diffusion establishes paths using *path reinforcement*; a node in the network may locally decide (based possibly on perceived traffic characteristics) to draw data from one or more neighbors in preference to other neighbors. In directed diffusion, exploratory samples have initially and repeatedly been flooded throughout the network. The sink then has some empirical information about which of its neighbors can provide it with the highest quality data (lowest loss or lowest delay). To this preferred neighbor, it sends out a reinforcement. That neighbor then locally determines its most preferred neighbor in the direction of the source, and so on. In our prior opportunistic approach, the most preferred neighbor is a neighbor which has delivered samples with the lowest delay.

In our greedy approach, each exploratory sample also contains an *energy cost* for delivering this sample from the source to the current node. In addition, each source on the existing tree (*i.e.,* a source on an established path) also generates an *on-tree incremental cost message* which corresponds to each new exploratory sample received. The incremental cost message contains the incremental energy cost required for delivering the corresponding exploratory sample to the existing tree. This incremental cost message is only sent and updated along the aggregation tree toward the sink. To find the closest point on the tree, the incremental energy-cost field can be updated only by closer nodes (*i.e.,*

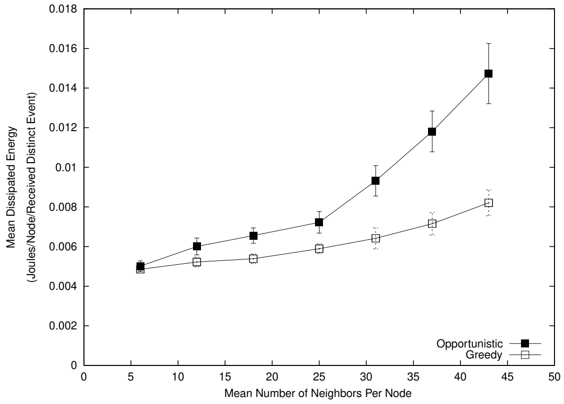

Figure 1. Greedy vs Opportunistic.

nodes which have received the corresponding exploratory sample at lower cost). In this greedy approach, the most preferred neighbor to reinforce is a neighbor which has delivered the exploratory sample or its corresponding incremental cost message at the lowest energy cost.

Data Aggregation and Set-Covering Problem: Intermediate nodes process or delay received data for a period of time before aggregating multiple messages into an aggregate. Before sending the aggregate, the nodes also compute associated energy cost for that aggregate. (This energy information will be used for path pruning.) However, different neighbors might report aggregates of different subsets of data items, with varying costs. The challenge is to find the set of incoming aggregates which cover the data items at the smallest cost. This is an instance of the weighted set-covering problem, for which we use heuristics from [1].

Path Pruning: Due to network dynamics, there can be multiple paths being reinforced. For energy efficiency, we need to prune unnecessary or inefficient paths. Our pruning rule is to *negatively reinforce* neighbors from which no energy-efficient aggregates have been received within a window of events or time. An incoming aggregate is considered energy-efficient if it is selected as a subset in the set cover. Specifically, we compute the set cover of sources that generate data items contained in received aggregates rather than the set cover of those data items.

3 Performance Evaluation

Figure 1 shows the mean dissipated energy per packet as a function of the network density, evaluated through simulation (see [4] for experiment details). Greedy aggregation has noticeably better energy efficiency than the opportunistic approach for very high-density networks. For dense sensor fields (40 or more neighbors per node), the dissipated energy is only 55% that of opportunistic aggregation. The reason for this benefit is that, given a high-density network, there exist several shortest paths from a source to a sink. The available path diversity reduces the probability of path sharing among different sources for opportunistic aggregation. Moreover, greedy aggregation has a delay and a distinct-event delivery ratio comparable to opportunistic aggregation (not shown here).

We also study the sensitivity of our greedy aggregation to several factors: network dynamics, number of sources, number of sinks, source placement schemes, and aggregation functions. These factors can adversely impact the energy savings of our approach, particularly when sources are near to one another and data is not perfectly aggregatable.

4 Conclusions

Our greedy approach constructs an energy-efficient aggregation tree using data-centric reinforcement mechanisms and prunes inefficient paths using a greedy heuristic for weighted set-covering problems. Our preliminary results suggest that although greedy aggregation and opportunistic aggregation are roughly equivalent in low-density networks, the greedy-tree approach can achieve signficant energy savings at high densities. In one experiment we find that greedy aggregation can achieve up to 45% energy savings over opportunistic aggregation without adversely impacting latency or robustness. Given that energy is scarce, this path optimization technique is essential for prolonging lifetime of highly-dense sensor networks.

References

[1] Thomas Bäck, Martin Schütz, and Sami Khuri. A comparative study of a penalty function, a repair heuristic, and stochastic operators with the set-covering problem. In *Artificial Evolution*, pages 3–20. Springer, Berlin, 1996.

[2] Deborah Estrin, Ramesh Govindan, John Heidemann, and Satish Kumar. Next century challenges: Scalable coordination in sensor networks. In *Proceedings of the Fifth Annual ACM/IEEE International Conference on Mobile Computing and Networking (Mobicom'99)*, Seattle, Washington, August 1999.

[3] John Heidemann, Fabio Silva, Chalermek Intanagonwiwat, Ramesh Govindan, Deborah Estrin, and Deepak Ganesan. Building efficient wireless sensor networks with low-level naming. In *Proceedings of the ACM Symposium on Operating Systems Principles*, Banff, Canada, October 2001.

[4] Chalermek Intanagonwiwat, Deborah Estrin, Ramesh Govindan, and John Heidemann. Impact of network density on data aggregation in wireless sensor networks. Technical Report 01-750, University of Southern California, November 2001.

[5] Chalermek Intanagonwiwat, Ramesh Govindan, and Deborah Estrin. Directed diffusion: A scalable and robust communication paradigm for sensor networks. In *Proceedings of the Sixth Annual ACM/IEEE International Conference on Mobile Computing and Networking (Mobicom'2000)*, Boston, Massachusetts, August 2000.

[6] Bhaskar Krishnamachari, Deborah Estrin, and Stephen Wicker. Impact of data aggregation in wireless sensor networks. In *International Workshop on Distributed Event-Based Sytems*, Vienna, Austria, July 2002.

[7] Gregory J. Pottie and William J. Kaiser. Embedding the internet: wireless integrated network sensors. *Communications of the ACM*, 43(5):51–58, May 2000.