# Evaluating Resource Tradeoffs and Optimizations in Data Dissemination Protocols for Wireless Sensor Networks

Amogh Kavimandan and Aniruddha Gokhale
Department of Electrical Engineering & Computer Science
Vanderbilt University
Nashville, TN 37235
Email: {amoghk,gokhale}@dre.vanderbilt.edu

**Abstract**

Performance of wireless sensor networks (WSNs) is dictated by their parameters, such as number of sensor nodes, sensor mobility and the network protocol used for routing. This paper provides two contributions to the study of data dissemination protocols in WSNs. First, we describe results investigating the performance of two data dissemination protocols – Directed Diffusion (DD) and Two-tier Data Dissemination (TTDD) – under varying WSN parameters and application properties. We compare their performance along two dimensions including (1) observed application latency and routing overhead, and (2) sensor network efficiency in terms of packet transmission ratio and energy consumption. Second, we describe our improvements to these protocols and illustrate the performance gains achieved by our changes.

Our experiments analyzing network performance indicate that increasing the number of nodes per unit sensing area does not make the network more redundant. At increased network densities and sink sizes the energy consumption starts increasing very rapidly on account of the added traffic. This argues for a better energy utilization mechanism, which scales the network as network density, while maintaining other performance parameters, such as delivery fraction and average delay, within acceptable limits. Our proposed protocol modifications call for local broadcast in DD and dynamically changing the flooding area in TTDD without changing the cell size.
**Keywords:** Wireless sensor networks, Data dissemination protocols, performance.

# Evaluating Resource Tradeoffs and Optimizations in Data Dissemination Protocols for Wireless Sensor Networks

## I. Introduction

A wireless sensor network (WSN) consists of low power, low cost, multifunctional and miniature collaborating sensor nodes, which are spread across a geographical area. The nodes are usually deployed inside or very close to a phenomenon being tracked. The sensor nodes carry a small processor, which can perform several localized computations, such as data aggregation or routing data packets. This enables them to process raw data generated by the phenomenon, filter it and send only the required information to the observer(s).

Moreover, the position of sensor nodes need not be predetermined thereby allowing any random deployment of these nodes and facilitating tracking of mobile phenomenon. These properties make the sensor networks applicable in a variety of applications ranging from military, to industrial, to scientific, to healthcare, to home. Most of these applications are aimed at (a) minimizing humans in the loop so that tasks can be executed remotely, such as remotely monitoring a patient's health and (b) collection of data in the environments that are hazardous to humans.

In WSNs a large number of sensor nodes collect the sensed data, which is then routed to a number of observers interested in this data. We call the sensor node generating data reports of an observed phenomenon as a *source node* and the node requesting the report as a *sink node*, both of which may be fixed or mobile. For example, soldiers collecting location information of the enemy using the sensor networks deployed in the battlefield act as sink nodes. The movement of the enemy acts as the phenomenon being tracked, whereas the sensor nodes in the immediate vicinity of that phenomenon act as source nodes. The source nodes detect the movement, generate appropriate data reports, which are then forwarded to each of the sink nodes interested in the event.

The unique characteristics of wireless sensor networks described below, however, makes the problem of disseminating the data generated by sensor nodes non-trivial and different than conventional wired networks. The main challenges are:

1) Wireless sensor networks are usually large to ensure that a mobile phenomenon can be covered, which might occur over a wide geographical terrain.
2) Computational power, storage capacity, battery life are limited, since the nodes themselves have to be large. Further, the nodes themselves can not be reconfigured (to add some resources for example) once deployed.
3) Node failures and frequent topology changes, due to node failures or sensor movement.

These problems form the motivation for a data-centric family of sensor network routing protocols called data-dissemination protocols, where the focus is on the data rather than the nodes. It follows that the protocols must be scalable, efficient, fault-tolerant and adaptable. A number of data-dissemination protocols have been suggested [10], [6], [4], [9], which aim to address these challenges.

Although significant work in WSN area has been done on designing and building sensor networks, most of it focuses on designing new applications and routing protocols. For example, Bokareva et al [2] recently proposed a comparative study of data-dissemination protocols comprising uniform node distribution under two different topologies. However, this is a more direct comparative analysis and does not study the effect of variations in availability of network resources on the overall performance. This paper aims to address the resource tradeoffs in designing a WSN and its ramifications on network performance and its design. We also compare two dissemination protocols *e.g.*, the Two-tier Data Dissemination (TTDD) [9] and Directed Diffusion (DD) [4] protocols, in order to provide useful insights into the design choices and optimizations for WSNs. The direct impact of our work is thus on providing useful insights in network and protocol optimizations, studying the best possible design and optimal resource decisions in the context of the dissemination protocols studied.

Our experiments analyzing network reliability indicate that increasing the number of nodes per unit sensing area does not make network more redundant; instead this might very well result in fast depleting network which is counter intuitive. One of the important findings from the analysis is that the the network needs to intelligently manage the resources. At increased network densities and sink sizes the energy consumption starts increasing very rapidly on account of the added traffic. This argues for a better energy utilization mechanism which scales the network well as network density, while maintaining other performance parameters (for example, delivery fraction, average delay) within acceptable limits.

**Organization:** The rest of the paper is organized as follows: Section II provides a brief overview of the data dissemination protocols used in our study; Section III describes the experimental setup used for our protocol simulations; Section IV presents and analyzes the results of our simulation studies; and finally Section V provides concluding remarks outlining lessons learned.

## II. Overview of Data Dissemination Protocols

This section provides a brief overview of the two data dissemination protocols we study in this paper.

**Directed Diffusion (DD):** Directed diffusion (DD) [4] was the first protocol to address the problem of sink mobility in the context of sensor networks. In DD each participating node maintains a dynamic map of the sensing tasks and the previous hop from which the *interest* was received. Each sensing task has an associated information, such as the node(s) to which to forward data reports, the rate of sending this data and the duration. As soon as the source nodes detect sensing tasks, they start sending exploratory packets to the sink nodes (which generate these tasks or interests). The sink node chooses one particular node (in its neighborhood) from which to receive the data and thus for a given source node, a single path is reinforced, which is used for all the traffic towards the sink.

**Two-tier Data Dissemination (TTDD):** Two-tier Data Dissemination (TTDD) [9] uses a two-tier architecture to disseminate the data to interested nodes. A grid is formed with the source node at one of the crossing points of the grid. Thus, the sensing space is divided by the grid into small cells. Only the nodes at the crossing points (called *dissemination nodes*) can forward the data. A sink node, which is interested in receiving a specific data, floods its local cell in order to discover the closest dissemination node. Each dissemination node contains forwarding information for each request of data it receives. On discovery of a dissemination point, the data request is forwarded to source node by the dissemination nodes, followed by data report being sent to the sink. Here, the smaller cell is the first tier, while the higher tier is made of the dissemination nodes at the grid points, which makes the two tier architecture.

## III. Evaluation Methodology

For studying the relationships of infrastructure and application properties of sensor networks, we used an evaluation environment within the NS-2 [7] discrete event simulator. The underlying MAC layer was IEEE 802.11, the physical radio characteristics of each node's network interface, such as transmit power, receiver sensitivity, data load and antenna type were chosen so as to approximate the Lucent WaveLan direct sequence spread spectrum radio. The evaluations were based on the simulation of sensor nodes which were distributed and moving within a square flat space of 1000m X 1000m for 100 seconds of simulation time. The mobility model used was the random waypoint model [5], [1].

The network parameters of interest to us was the sensor network size. This parameter were varied for the different protocols described in Section II. Increasing the total nodes in a network will not necessarily increase the network performance, for the same topology since denser node will have more collisions, possibly node failures and hence congestion. The simulations are run for 100, 196, 289 and 400 nodes for the same square flat space described above [8].

The data-dissemination protocols mainly address the observer mobility issue. As such it is critical to study the behavior of protocols varying the following parameters:

1) **number of observers (*i.e.*, sinks nodes):** Network performance against increasing sink nodes is an important metric since it is indicative of how well the network scales as sink nodes. We call the total number of sink nodes in the simulation as sink size.
2) **pause time:** In the sensor network, each node remains initially stationary for a fixed period of time, known as the pause time. A destination is then selected from the flat simulation space defined and then a node starts moving towards the destination. The speed of the node varies from 0 to some maximum speed specified at the beginning of simulation. On reaching the destination, the node again pauses for the fixed time, at the end of which it selects another destination and the process is repeated for simulation time. Pause time is thus a measure of sink mobility, since lower pause time indicates highly mobile sinks, while longer pause times mean the sinks are stationary for longer periods. Increased sink movements leads to more frequent changes in dissemination path (and potentially increased traffic). Thus it is interesting to observe protocol efficiency for increased sink mobility. Such an experimental setup was important in order to challenge the protocols in an environment where they would be most applicable.

The simulations are run for reports generated (by source nodes) every 0.25 seconds. In order to have a direct comparison for the protocols being studied, the same environment variables with identical loads were used for the all experiments. The same movement files, which describe the movement of nodes in the sensor network are used for both the protocols. The simulations were run with movement patterns which were generated for seven different pause times including 1, 2, 10, 25, 40, 50 and 75 seconds. Eight different values for the number of sink nodes were used including 2, 4, 6, 8, 10, 12, 14 and 16 sinks nodes.

## IV. Results and Discussion

This section describes results and analysis of our experiments described in Section III. The results are presented along different dimensions of WSN performance metrics.

### A. Aggregate Energy Consumption

Aggregate Energy Consumption is defined as the total energy consumed (in joules) in transmission and receiving, during the simulation. The Sink Size is the total number of sink nodes in the wireless network. Assume that the total energy per node $i$ is $E_i$ and energy consumption for this node is $E_{consumed_i}$, then the energy depletion, $E_{depletion}$, is defined as

$$E_{depletion} = \sum_{i=1}^{n}(E_i - E_{consumed_i})/n \qquad (1)$$

where $n$ is the network density.

Figures 1 and 2 show the variation of aggregate energy consumption against sink size for various network densities for DD and TTDD, respectively. On an average, the energy consumption increases by 17% for a 60% increase in the sensor nodes deployed. It would appear that increasing the total sensors per unit flat space under observation will increase the redundancy of the network. This follows from the fact that since there are more nodes per unit space, at least some nodes will always be "alive" and would route the through traffic for that space. However, from Figures 1 and 2, we observe that the energy gets depleted faster for highly dense nodes. For example, for the same sink size (of 2) the energy depletion is 36% for a network density of 100, while it is 50% for a network density of 400. Thus, a highly dense network will not increase the redundancy. This is because for such a network for the same time the energy of a single node will deplete at a faster rate and hence the nodes that are "alive" per unit flat space will be lesser.

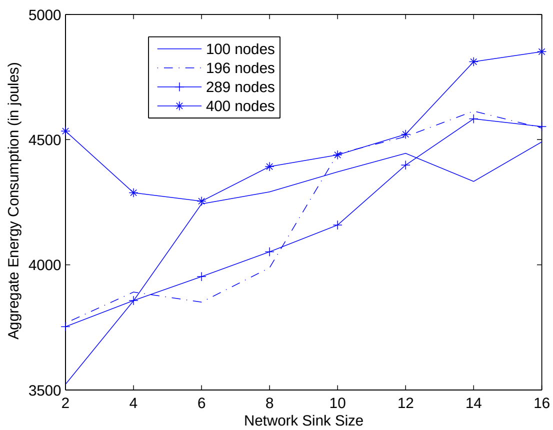


Fig. 1. **Average Energy Consumption During Transmit/Receive for DD**

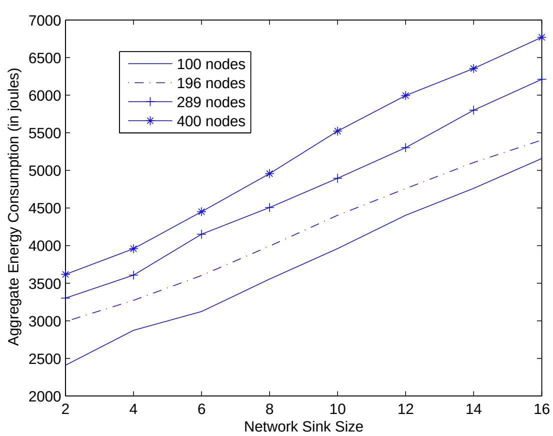


Fig. 2. **Average Energy Consumption During Transmit/Receive for TTDD**

## B. Packet Delivery Fraction discussions

Packet Delivery Fraction is defined as the fraction of number of packets received at the sink to the number of packets being sent from the source nodes. The fraction is averaged for all the sink-source pairs for each network density in the simulation. The variation of packet delivery fraction against sink size, for network densities of 100, 196, 289 and 400 for DD and TTDD, respectively, is shown in Figures 3 and 4.

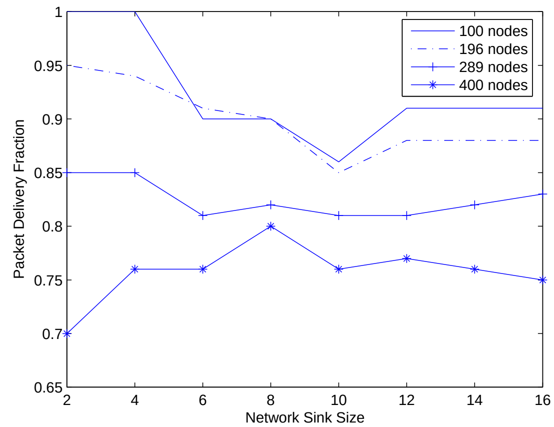


Fig. 3. **Packet Delivery Fraction versus Sink Size for DD**

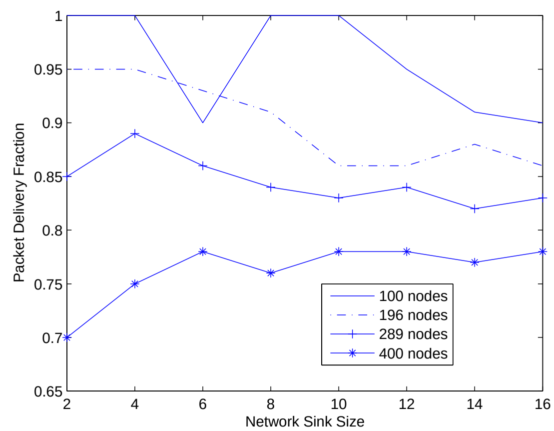


Fig. 4. **Packet Delivery Fraction versus Sink Size for TTDD**

We observe that the fraction variation is about 7% on an average, for increase in sink size from 2 to 16. This reduction is owing to an increase in the overall traffic since the reports have to reach more observers. This allows us to accommodate more sinks (an 800% increase) at the cost of 7% more packets getting dropped. This might seem a reasonable tradeoff, however, it must be noted that for the same sink size increase, the energy consumption increased almost 100%! A better energy utilization, in particular, as the sink size is increased is thus crucial to increase network performance, since this will mean more sinks could be tolerated at an acceptable energy consumption level and delivery fraction values.

## C. Normalized Routing Load discussions

Routing load (L) for a single node is defined as the total number of routing packets that it transmits in a network for the duration of simulation. We define Normalized Routing Load ($L_n$) as

$$L_n = (\sum_{i=0}^{n} L_i)/n \tag{2}$$

where $n$ is the total number of nodes in the network.

Figures 5 and 6 show the variation of normalized routing load against pause time for various network densities for DD and TTDD, respectively. The sink size is kept constant at 6 for this set of experiments. Observe that, for less mobile sinks (indicated by higher pause times), the TTDD has more normalized routing load than DD. This is expected since, TTDD has to construct and maintain the grid which is very costly. Further, TTDD has a localized flooding, where a small area of topology is flooded till a dissemination point is located. The flooding area is thus another factor which contributes the routing load. The flooding area is a static configuration parameter, and thus once set it can not be dynamically varied (in order to reduce the normalized routing load, since lesser are will mean lesser flooding). Thus, TTDD has more routing load than DD for the same pause time and sink size.

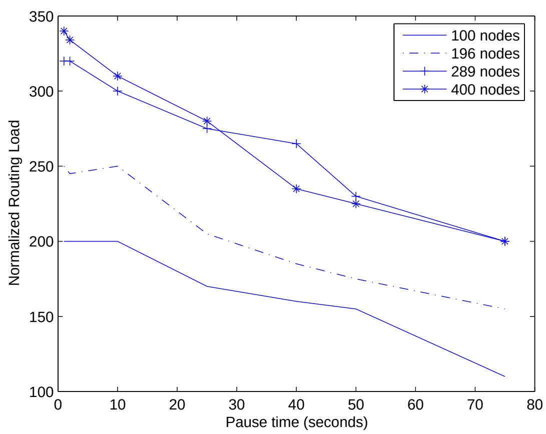


Fig. 5. **Normalized Routing Load versus Pause (in seconds) for DD**

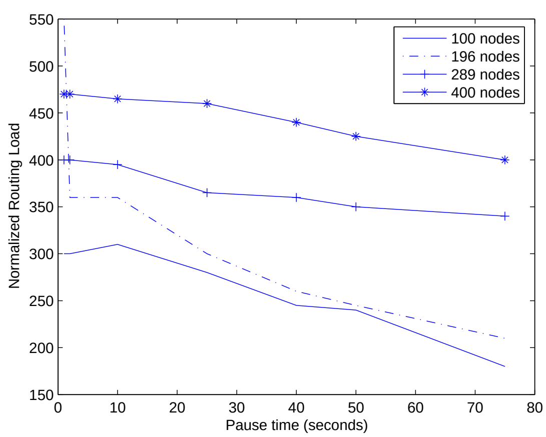


Fig. 6. **Normalized Routing Load versus Pause (in seconds) for TTDD**

## *D. Average Delay discussions*

The average delay is defined as the latency seen by the sink in obtaining reports of interests from the source, averaged for all sink-source pairs. Figures 7 and 8 show the variation against sink size for DD and TTDD, respectively. As one would expect, with the increase in sink size, more packets are transmitted for the same per unit sensing area causing increased delays as seen by the sink. Figures 7 and 8 confirm this behavior where the average delay increases steadily as sink size increases. One important observation from the TTDD and DD delay results is that the delay increases more rapidly for TTDD than for DD. This is corroborated by the fact that other parameters being constant, the routing overhead is more for TTDD than DD, which leads to congestion and hence an increase in the delays.

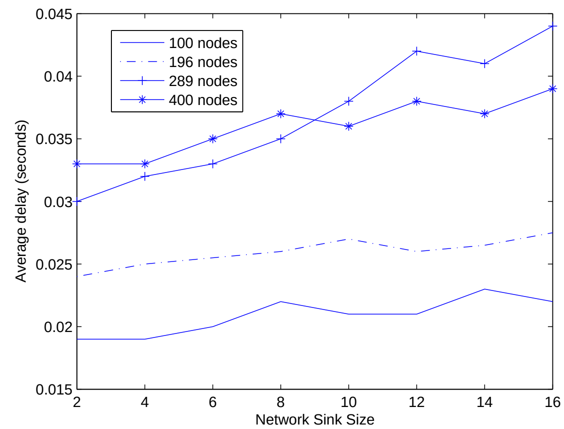


Fig. 7. **Average Delay (in seconds) versus Sink Size for DD**

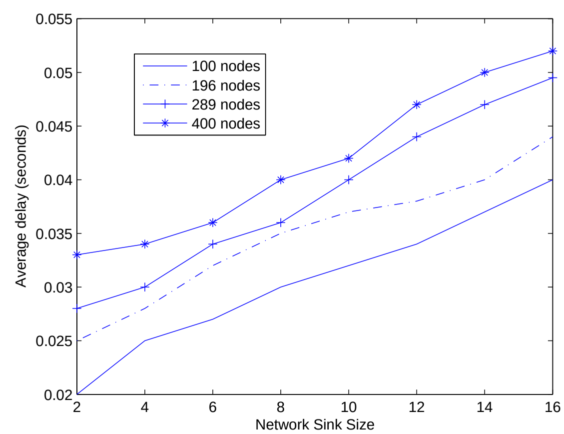


Fig. 8. **Average Delay (in seconds) versus Sink Size for TTDD**

## *E. Suggestions on protocol improvement*

Based on insights gained from our experimental evaluation, we now briefly describe proposed improvements to the two protocols we studied.

1) **DD:** In DD, network-level broadcast (for advertising interest in specific sensing data) and sending exploratory data packets are one of the most expensive operations and hence network performance will be greatly affected if the cost associated with these operations is reduced.

One possible solution is diffusing interests locally instead of broadcast, such as using passive clustering [3], where the redundant transmissions during broadcast operations are reduced effectively reducing the costs. Furthermore, cost reduction can be done by exploring multiple paths instead of a single reinforced path since a single path depletes the nodes that make up this path faster.

2) **TTDD:** In TTDD, the flooding is restricted to a single home cell of a sink node, thus aiming to reduce the associated cost. However, the flooding area is not dynamic and thus once set, the same area will be flooded with sensing data request. A possible enhancement to TTDD can be that if the current density of the network and average energy consumption makes cell-level broadcast costly, a new tier could be introduced at the cell-level. A new dissemination point can be appointed inside the cells themselves. This is better than changing the cell size [2] because changing cell-size dynamically means reconstructing the grid structure, which is a very costly operation. In the worst case scenario, for a sink lying on the grid but at a distance $\alpha/2$ from each of the closest dissemination points, the flooding will be done for a maximum of 0.785 unit area for unit $\alpha$. The worst case value for the plain two grid structure without the suggested improvement is 1.57 unit area, thus giving a reduction of almost 50%.

## V. Concluding Remarks

In this paper we discussed the resource tradeoffs for WSNs. We studied the protocol performance against steady increase of sensor nodes and sink size in the network. The contributions and lessons learned from our work are twofold.

**1. Effects of resource tradeoffs in sensor networks:** We studied the effects of resource tradeoffs on network performance, which indicated the following:

1) We show that increasing nodes per unit sensing area does not make the network mode redundant, instead this may lead to a fast depleting network which is counter productive.
2) At increased network densities and sink sizes (number of sinks in the network), the energy consumption starts increasing very rapidly due to increased traffic that each node has to handle. This argues for an intelligent energy utilization mechanism which scales the network well as the above parameters.

**2. Proposed enhancements to protocols:** Based on our analysis, we propose the following improvements to the two protocols studied.

1) In Directed Diffusion (DD), we propose a local broadcast, which reduces redundant transmissions during broadcast operations. Further, performance improvement can be achieved by maintaining multiple paths instead of a single reinforced path, to reduce the energy depletion of the single path nodes.
2) In Two-tier Data Dissemination (TTDD), we propose a new mechanism to dynamically change the flooding area without changing the cell size *i.e.*, without having to reconstruct the grid. In the worst case, our approach gives a reduction of almost 50% in the local (cell) flooding area.

## References


[1] Christian Bettstetter, Giovanni Resta, and Paolo Santi. The node distribution of the random waypoint mobility model for wireless ad hoc networks. *IEEE Transactions on Mobile Computing*, 2(3):257–269, Jul-Sep 2003.

[2] Tatiana Bokareva, Nirupama Bulusu, and Sanjay Jha. A performance comparison of data dissemination protocols for wireless sensor networks. In *Proceedings of the Forty Seventh Annual IEEE Global Telecommunications Conference (GLOBECOM 2004) Ad hoc and Wireless Sensor Networks Track*, Dallas, TX, 2004.

[3] Vlado Handziski, Andreas Kpke, Holger Karl, Christian Frank, and Witold Drytkiewicz. Improving the energy efficiency of directed diffusion using passive clustering. In *Proceedings of the 1st European Workshop on Wireless Sensor Networks (EWSN)*, pages 172–187, Berlin, Germany, January 2004. Springer-Verlag, ISBN 3-540-20825-9.

[4] 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)*, pages 56–67, Boston, MA, 2000.

[5] David B Johnson and David A Maltz. Dynamic source routing in ad hoc wireless networks. In Imielinski and Korth, editors, *Mobile Computing*, volume 353. Kluwer Academic Publishers, 1996.

[6] Gunjan Khanna, Saurabh Bagchi, and Yu-Sung Wu. Fault tolerant energy aware data dissemination protocol in sensor network. In *Proceedings of the IEEE Dependable Systems and Networks Conference (DSN 2004)*, Florence, Italy, June 2004.

[7] The VINT Project. Network Simulator - NS-2. `http://www.isi.edu/nsnam/ns`.

[8] Sameer Tilak, Nael B. Abu-Ghazaleh, and Wendi Rabiner Heinzelman. Infrastructure tradeoffs for sensor networks. In *Proceedings of the First International Workshop on Wireless Sensor Networks and Applications (WSNA 2002)*, Atlanta, GA, September 2002.

[9] Fan Ye, Haiyun Luo, Jerry Cheng, Songwu Lu, and Lixia Zhang. Two-tier data dissemination model for large-scale wireless sensors networks. In *Proceedings of the Eight Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom 2002)*, pages 148–159, Atlanta, GA, 2002.

[10] Fan Ye, Haiyun Luo, Songwu Lu, and Lixia Zhang. Gradient broadcast: A robust data delivery protocol for large-scale sensor networks. In *Proceedings of International Conference on Information Processing in Sensor Networks (IPSN 03)*, Palo Alto, CA, April 2003.