Power Efficient Data Gathering and Aggregation in Wireless Sensor Networks

Hüseyin Özgür Tan and İbrahim Körpeoğlu Department of Computer Engineering, Bilkent University 06800 Ankara, Turkey E-mail:{hozgur,korpe}@cs.bilkent.edu.tr

Abstract-Recent developments in processor, memory and radio technology have enabled wireless sensor networks which are deployed to collect useful information from an area of interest. The sensed data must be gathered and transmitted to a base station where it is further processed for end-user queries. Since the network consists of low-cost nodes with limited battery power, power efficient methods must be employed for data gathering and aggregation in order to achieve long network lifetimes.

In an environment where in a round of communication each of the sensor nodes has data to send to a base station, it is important to minimize the total energy consumed by the system in a round so that the system lifetime is maximized. With the use of data fusion and aggregation techniques, while minimizing the total energy per round, if power consumption per node can be balanced as well, a near optimal data gathering and routing scheme can be achieved in terms of network lifetime.

So far, besides the conventional protocol of direct transmission, two elegant protocols called LEACH and PEGASIS have been proposed to maximize the lifetime of a sensor network. In this paper, we propose two new algorithms under name PEDAP (Power Efficient Data gathering and Aggregation Protocol), which are near optimal minimum spanning tree based routing schemes, where one of them is the power-aware version of the other. Our simulation results show that our algorithms perform well both in systems where base station is far away from and where it is in the center of the field. PEDAP achieves between 4x to 20x improvement in network lifetime compared with LEACH, and about three times improvement compared with PEGASIS.

I. Introduction

With the introduction of low-cost processor, memory, and radio technologies, it becomes possible to build inexpensive wireless micro-sensor nodes. Although these sensors are not so powerful compared to their expensive macro-sensor counterparts, by using hundreds or thousands of them it is possible to build a high quality, fault-tolerant sensor network. These networks can be used to collect useful information from an area of interest, especially where the physical environment is so harsh that the macro-sensor counterparts cannot be deployed. They have a wide range of applications, from military to civil, that may be realized by using different type of sensor devices with different capabilities for different kinds of environments [1].

The main constraint of sensor nodes is their very low finite battery energy, which limits the lifetime and the quality of the network. For that reason, the protocols running on sensor going research on power management issues in order to reduce the power consumption when the nodes become idle [2]. When power efficient communication is considered, it is important to

networks must consume the resources of the nodes efficiently in order to achieve a longer network lifetime. There is an on-

SIGMOD Record, Vol. 32, No. 4, December 2003

maximize the nodes' lifetimes, reduce bandwidth requirements by using local collaboration among the nodes, and tolerate node failures, besides delivering the data efficiently.

There are several power efficient protocols defined for wireless ad-hoc networks ([3], [4]). When sensor networks are considered, Chang and Tassiulas ([5], [6]) give data routing algorithms which maximizes the system lifetime where only some of the nodes have data to send and where there can be more than one base station. In another work by Bhardwaj et al. [7], the upper bounds on the lifetime of sensor networks are derived. There are also different protocols proposed in the literature ([8], [9], [10]) to maximize the lifetime of the system under different circumstances.

Since data generated in a sensor network is too much for an end-user to process, methods for combining data into a small set of meaningful information is required. A simple way of doing that is aggregating (sum, average, min, max, count) the data originating from different nodes. A more elegant solution is data fusion which can be defined as combination of several unreliable data measurements to produce a more accurate signal by enhancing the common signal and reducing the uncorrelated noise [11]. These approaches have been used by different protocols ([11], [12]) so far, because of the fact that they improve the performance of a sensor network in an order of magnitude by reducing the amount of data transmitted in the system.

There are various models for sensor networks. In this work we mainly consider a sensor network environment where:

- Each node periodically senses its nearby environment and would like to send its data to a base station located at a fixed point.
- Sensor nodes are homogeneous and energy constrained.
- Sensor nodes and base station are stationary.
- Data fusion or aggregation is used to reduce the number of messages in the network. We assume that combining n packets of size k results in one packet of size k instead of size nk.

The aim is efficient transmission of all the data to the base station so that the lifetime of the network is maximized in terms of rounds, where a round is defined as the process of gathering all the data from sensor nodes to the base station, regardless of how much time it takes.

Direct transmission is a simple approach for this problem in which each node transmits its own data directly to the base station. However, if the base station is far away, the cost of sending data to it become too large and the nodes will die quickly. In order to solve this problem, two elegant protocols LEACH [11] and PEGASIS [12] are proposed. In LEACH, the key idea is to reduce the number of nodes communicating directly with the base station. The protocol achieves this by forming a small number of clusters in a selforganizing manner, where each cluster-head collects the data from nodes in its cluster, fuses and sends the result to the base station. LEACH also uses randomization in cluster-head selection and achieves up to 8x improvement compared to the direct transmission approach. PEGASIS takes it further and reduces the number of nodes communicating directly with the base station to one by forming a chain passing through all nodes where each node receives from and transmits to the closest possible neighbor. The data is collected starting from each endpoint of the chain until the randomized head-node is reached. The data is fused each time it moves from node to node. The designated head-node is responsible for transmitting the final data to the base station. PEGASIS achieves a better performance than LEACH by between 100% and 300% in terms of network lifetime.

In this work, we propose a new minimum spanning tree-based protocol called PEDAP (Power Efficient Data gathering and Aggregation Protocol) and its power-aware version. PEDAP prolongs the lifetime of the last node in the system while providing a good lifetime for the first node, whereas its power-aware version provides near optimal lifetime for the first node although slightly decreasing the lifetime of the last node. Another advantage of our protocols is they improve the lifetime of the system even if the base station is inside the field, whereas LEACH and PEGASIS cannot.

The rest of the paper is organized as follows. In Section 2, we formulate our system model and the data gathering problem. The PEDAP protocols are described in detail in Section 3. Next, in Section 4 we present our simulation results compared with other known algorithms. The feasibility of implementation of our algorithms is discussed in Section 5. Finally, we conclude the paper and present future research directions in Section 6.

II. SYSTEM MODEL AND PROBLEM

A. Radio Model

Recently, there is a significant amount of work in the area of building low-energy radios. In our work, we used the first order radio model presented in [11]. In the specified radio model, the energy dissipation of the radio in order to run the transmitter or receiver circuitry is equal to $E_{elec}=50nJ/bit$, and to run the transmit amplifier it is equal to $E_{amp}=100pJ/bit/m^2$. It is also assumed an r^2 energy loss due to channel transmission. Therefore, the energy expended to transmit a k-bit packet to a distance d and to receive that packet with this radio model is:

$$E_{Tx}(k,d) = E_{elec} * k + E_{amp} * k * d^2$$
 (1)

$$E_{Rx}(k) = E_{elec} * k (2)$$

It is also assumed that the radio channel is symmetric, which means the cost of transmitting a message from A to B is the same as the cost of transmitting a message from B to A.

As mentioned in [11], the energy required for receiving a message is not so low. Therefore, the routing protocols must also minimize the number of receive and transmit operations for a specific node while minimizing the transmit distances.

It is also important to note that the cost of one transmission of a k-bit packet to the system is either:

$$C_{ij}(k) = 2 * E_{elec} * k + E_{amp} * k * d_{ij}^2$$
 (3)

٥r

$$C'_{i}(k) = E_{elec} * k + E_{amp} * k * d_{ib}^{2}$$
 (4)

where C_{ij} is the cost of transmission between node i and node j, C_i' is the cost between node i and the base station, d_{ij} is the distance between node i and node j, and d_{ib} is the distance between node i and the base station. Since C_i' is smaller than C_{ij} when the term with E_{amp} is much smaller than the term with E_{elec} , for the overall system lifetime it can be advantageous to increase the number of transmissions to the base station.

The parameter values used in our work are the same as those used in LEACH and PEGASIS, in order to see the level of energy savings that our protocols can achieve.

B. Problem Statement

In this work, our main consideration is wireless sensor networks where the sensors are randomly distributed over an area of interest. The locations of sensors are fixed and the base station knows them all a priori*. The sensors are in direct communication range of each other and can transmit to and receive from the base station. The nodes periodically sense the environment and have always data to send in each round (period) of communication. The nodes aggregate or fuse the data they receive from the others with their own data, and produce only one packet regardless of how many packets they receive.

The problem is to find a routing scheme to deliver data packets collected from sensor nodes to the base station, which maximizes the lifetime of the sensor network under the system model given above. However, the definition of the lifetime is not clear unless the kind of service the sensor network provides is given. In applications where the time that all the nodes operate together is important, - since the quality of the system will be dramatically decreased after first node death lifetime is defined as the number of rounds until the first sensor is drained of its energy. In another case, where the nodes are densely deployed, the quality of the system is not affected until a significant amount of nodes die, since adjacent nodes record identical or related data. In this case, the lifetime of the network is the time elapsed until half of the nodes or some specified portion of the nodes die. In general, the time in rounds where the last node depletes all of its energy defines the lifetime of the overall sensor network. Taking these different possible requirements under consideration, our work

^{*}This information can be entered manually to the base station, or the base station can get the coordinates from the nodes if the nodes are equipped with GPS, or alternatively techniques like triangulation can be used.

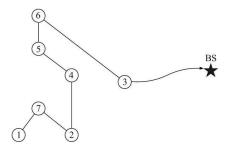


Fig. 1. Chain based routing scheme on a sample network.

gives timings of all deaths for all algorithms in detail and leaves the decision which one to choose to system designers.

C. Energy Analysis for Data Routing

In [11], the energy dissipations in MTE (Minimum-Transmission-Energy) routing and direct transmission are compared and it is figured out that an ideal system must use a hybrid of both when the base station is far away from the nodes. The authors propose a two-level clustering hierarchy based routing scheme, in which the number of nodes (clusterheads) that transmit data to the base station is reduced to 5%, while all of other nodes determine their closest gateway (cluster-head) to the base station in order to send their data. The cluster-heads are chosen randomly in order to make the system lifetime longer. However, since this algorithm is purely random, it is far from optimal.

In [12], authors noticed that in a close neighborhood, the cost of running receive or transmit circuitry is larger than the cost of running the amplifier circuitry for a single node. So they propose a scheme where all nodes receive and transmit only once over the edges of a chain passing through all nodes and whose length is close to minimum. In each round, a special node is selected randomly to send the fused data to the base station. Thus, only one node communicates with the base station. The algorithm works fine when the base station is far away from the field in which case the cost of sending data to the base station is almost the same for all nodes. In that case, regardless of who sends data to the base station, for a round of communication the algorithm tries to minimize the energy consumed by each node, in turn maximizes the lifetime of the nodes. Figure 1 shows a routing scheme that [12] computes for a sample network.

However, when the base station is inside the field (close to the center), both of the protocols perform poor. This is mainly because they do not take the exact cost of sending data to base station into account and make a decision according to that. In addition to this, the approaches so far have not considered minimizing the total energy consumed per-round in the system.

We believe that the main idea, in order to maximize the network lifetime, should be to minimize the total energy expended in the system in a round of communication, while balancing the energy consumption among the nodes.

The first part of the idea can be realized optimally by computing a minimum spanning tree over the sensor network SIGMOD Record, Vol. 32, No. 4, December 2003

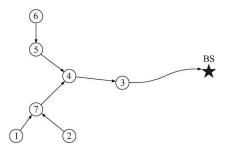


Fig. 2. Minimum spanning tree based routing scheme on a sample network.

with link costs C_{ij} (given in Equation 3) among the nodes and C'_i (given in Equation 4) between the nodes and base station. The data packets are then routed to the base station over the edges of the computed minimum spanning tree. We call this routing strategy as PEDAP (Power Efficient Data gathering and Aggregation Protocol). Figure 2 illustrates the idea on a sample network. Although PEDAP does not take the balancing issue into account, it always achieves a good lifetime for the last node. This is because, until the time the first node dies, the minimum possible energy is expended from the whole system. So the total remaining energy is optimum for the rest of the nodes. This is true for each death, thus after each node death the remaining energy in the system is maximum. So PEDAP protocol achieves almost the optimum lifetime for the last node in the system, while providing a good lifetime for the first node.

In order to achieve the second part of the idea, balancing the load (henceforth the energy consumption) among the nodes, we can use the information about the remaining energy of each node. When the base station is far away from the nodes, the node that dies first is usually the one that sends aggregated and fused data to the base station. So, a node with low remaining energy would not want to send to the base station. That node would like to expend its remaining energy by sending to a nearby neighbor and thus try to maximize its lifetime. Also a low-energy node would not like to receive many packets from others, since receiving is a high cost operation too. Its tendency would be only to send its data and not to receive anything from others. In order to achieve these, a slight change in the cost functions helps us. The new cost functions will be as follows:

$$C_{ij}(k) = \frac{2 * E_{elec} * k + E_{amp} * k * d_{ij}^{2}}{e_{i}}$$
(5)
$$C'_{i}(k) = \frac{E_{elec} * k + E_{amp} * k * d_{ib}^{2}}{e_{i}} ,$$
(6)

$$C'_{i}(k) = \frac{E_{elec} * k + E_{amp} * k * d_{ib}^{2}}{e_{i}},$$
 (6)

where e_i is the remaining energy of node i, which is normalized with respect to the maximum possible energy in the battery (i.e. $0 \le e_i \le 1$).

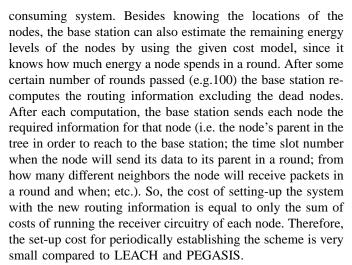
As it can be noticed, now the cost of communication between the nodes is not symmetric. According to Equation 5, the cost of sending a message from a node i to its neighbors increases as the remaining energy of node i decreases. Although this new formula usually does not change the selection of the neighbor which a node sends, it postpones the inclusion of that node in the spanning tree. The later a node is included in the spanning tree, the fewer number of messages it will receive. According to Equation 6, for a low-energy node the cost of sending to the base station is increased, and thereby the willingness to send to the base station for that node is decreased. So, if the minimum spanning tree algorithm would be executed periodically every certain number of rounds (such as 100), a more power efficient routing scheme is found for the next period, depending on the current situation (the nodes that are alive and their energy levels). This is the idea behind the power-aware version of PEDAP, which we will call PEDAP-PA (Power Efficient Data gathering and Aggregation Protocol - Power Aware).

III. PEDAP ALGORITHMS DETAILS

The PEDAP protocols assume the locations of all nodes are known by base station a priori. They are both centralized algorithms where the base station is responsible for computing the routing information. This is because, in systems where some elements are resource limited whereas one or more elements are powerful, it is desirable to give the computation load to the more powerful elements of the system.

The routing information is computed using Prim's minimum spanning tree algorithm where base station is the root. The algorithm works as follows: Initially, we put a node in the tree which is the base station in our case. After that, in each iteration we select the minimum weighted edge from a vertex in the tree to a vertex not in the tree, and add that edge to the tree. In our case this means that the vertex just included in the tree will send its data through that edge. We repeat this procedure until all nodes are added to the tree. In Figure 2, the resulting routing paths are illustrated for a sample network. The running time complexity of the algorithm is $O(n^2)$ assuming there are n nodes in the network.

As seen, the base station is included in the network graph. Thus, by computing a minimum spanning tree over this graph with the cost functions given as above and by routing packets according to that spanning tree, we achieve a minimum energy



For the two algorithms proposed in this work, the protocols are the same. Only thing that must be changed is the cost functions. So switching between the two proposed algorithms requires only a small change in the base station and no changes in sensor nodes. This makes our algorithms preferable when different applications with different lifetime requirements will be executed in the same sensor network from time to time.

IV. SIMULATION RESULTS

In order to evaluate the performance of our algorithms, we simulated five different routing schemes: Direct transmission, LEACH, PEGASIS, PEDAP and PEDAP-PA, the later two being our proposals. The simulations are done in C. We generate networks of diameters 50 m and 100 m randomly, each having 100 nodes. We repeated the simulations for the same network twice: one with a distant base station, other with a base station in the center. We located the base station to point (0, -100) in simulations where it is distant. We ran the simulations with different network sizes and different initial energy levels. The aim was to determine the timings of node deaths (in terms of rounds) until the last node dies. Once a node dies, we consider it dead for the rest of the

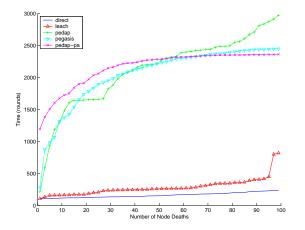


Fig. 3. Timings of node deaths in a network of size 50 m x 50 m - The base station is distant from the field.

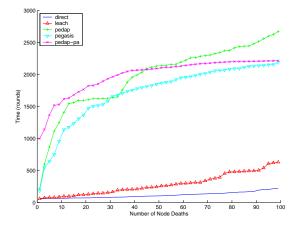


Fig. 4. Timings of node deaths in a network of size 100 m x 100 m - The base station is distant from the field.

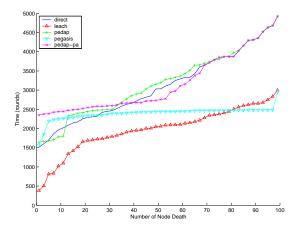


Fig. 5. Timings of node deaths in a network of size 50 m x 50 m - The base station is in the center.

simulation. We re-computed the routing information every 100 rounds for all algorithms. This parameter is important for the actual system performance. A small value leads better results for all algorithms. However, in that case the set-up costs, which are not included in the simulations, may dominate the communication costs.

Figure 3 and Figure 4 show the timings of all deaths for networks whose diameters are 50 m and 100 m respectively and where the base station is far away from the field. As seen, while LEACH and direct transmission perform far from optimal, PEGASIS provides a good improvement in both cases. However, PEDAP-PA further improves the lifetime of the first node about 400%, while providing almost the same lifetime for the last node, compared with PEGASIS. On the other hand, compared again with PEGASIS, PEDAP improves the lifetime of the last node about 25%, while providing almost the same lifetime for the first node.

TABLE I
TIMINGS OF NODE DEATHS. BASE STATION IS IN THE CENTER

Energy(J)	Protocol	FND	HND	LND
0.25	DIRECT	596	1147	4836
	LEACH	297	1247	2223
	PEGASIS	439	2259	2667
	PEDAP	1228	2334	4836
	PEDAP-PA	2177	2352	4836
0.50	DIRECT	1192	2293	9672
	LEACH	1036	2927	4362
	PEGASIS	774	4496	5175
	PEDAP	2455	4668	9672
	PEDAP-PA	4353	4688	9672
1.00	DIRECT	2383	4586	19343
	LEACH	2627	5603	7747
	PEGASIS	1428	9036	10443
	PEDAP	4910	9336	19343
	PEDAP-PA	8705	9378	19343

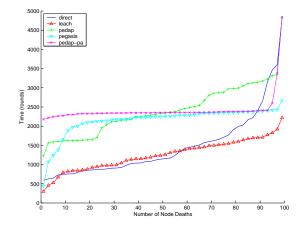


Fig. 6. Timings of node deaths in a network of size 100 m x 100 m - The base station is in the center.

Figure 5 and 6 provide the same information, but in this case the base station is located in the center of the field. Now, both PEDAP and PEDAP-PA improve the lifetime of the last node about two times compared with PEGASIS and LEACH. The improved value is the same as it is in direct transmission, which is optimum. As for the first node death time, both PEDAP protocols improve it when compared with others. However, PEDAP-PA achieves about two times improvement over PEDAP. Therefore, we can conclude that PEDAP-PA is the best performing algorithm for systems where base station is in the center of the field. It gives the best lifetime for the first node, while providing the optimum lifetime for the last. However, if the nodes are not power-aware, PEDAP is a good alternative for the same environment.

Table I and Table II summarize the results for two different base station locations and for three different initial energy levels in a network of diameter 100 m. In these tables, FND

 $\label{table II} \mbox{Table II}$ Timings of node deaths. Base station is distant from the field

Energy(J)	Protocol	FND	HND	LND
0.25	DIRECT	61	104	223
	LEACH	60	255	632
	PEGASIS	184	1856	2190
	PEDAP	213	2135	2674
	PEDAP-PA	998	2103	2217
0.50	DIRECT	121	208	445
	LEACH	123	661	2134
	PEGASIS	1070	3767	4344
	PEDAP	426	4271	5337
	PEDAP-PA	2897	4067	4272
1.00	DIRECT	242	416	889
	LEACH	351	1983	3961
	PEGASIS	1332	7309	8536
	PEDAP	851	8544	10665
	PEDAP-PA	6899	7763	8438

and LND stand for the times at which the first and the last node die. HND stands for the time at which half of the nodes die. Note that the performance of LEACH is very close to direct communication in all simulations. This is because we recompute the routing information every 100 rounds, which is a reasonable number. So, LEACH protocol consumes too much energy until 100^{th} round.

Also it is worth to note that doubling initial energy level almost doubles the lifetimes in all protocols as expected. In PEDAP-PA, however, when the initial energy is doubled, the lifetime of first node increases about 2.5 times. We believe that this is because PEDAP-PA finds more chance to recompute the routing information with increasing initial energy. As the number of re-computations increases, more energy saving routing paths are achieved, since PEDAP-PA is power-aware unlike others.

V. IMPLEMENTATION

In this section, we consider firstly the basic environment for implementing our algorithms in a real-life situation. After that, we discuss other environments where our algorithms are also feasible to implement. First of all, for our algorithms to work, every node should be able to communicate with base station and with each other. The locations of nodes must be known by the base station a priori, which can be done either by manually entering coordinates, or by means of methods such as triangulation, or by using GPS. We do not consider the length of a round. This is reasonable for applications where the measurements are taken infrequently such as periodic measurements of average temperature in an area of interest.

We divided each round into stages whose length is equal to the time to send a message multiplied by the maximum of the in-degrees of the nodes in the minimum spanning tree. The number of stages is determined by the depth of the tree. In the first stage, all leaf nodes at maximum depth send their data to their parents. The parents apply TDMA multiple access scheme among their children. Each node sends its message with its parents CDMA code, in order to prevent collisions with the messages of other nodes sending to different parents at the same time. In the next stages, the procedure climbs one level up until it reaches the root, the base station. After 100 such rounds all nodes stop sending their data, and turn on their receivers to get the information about the new routing paths computed by the base station.

Our algorithms can also work in environments where all the nodes and the base station are not in direct communication range of each other. In this case, a distributed minimum spanning tree algorithm [13] can work. However, this method increases set-up cost dramatically. On the other hand, if the base station can still transmit to all the nodes directly, the scheme can be efficiently computed at the base station assumming the visibility graph is given.

VI. CONCLUSION AND FUTURE WORK

In this work, we present PEDAP and PEDAP-PA, two power efficient data gathering and aggregation protocols based on

minimum spanning tree routing scheme. We show through simulations that our algorithms perform near optimal. PEDAP outperforms previous approaches, LEACH and PEGASIS, by constructing minimum energy consuming routing for each round of communication. PEDAP-PA takes it further and tries to balance the load among the nodes. Minimizing the total energy of the system while distributing the load evenly to the nodes has a great impact on system lifetime. This is confirmed through simulations.

Our simulations show that if keeping all the nodes working together is important, PEDAP-PA performs best among others, regardless of the position of the base station. On the other hand, if the lifetime of the last node is important or the nodes are not power-aware, PEDAP is a good alternative.

It is worth to note that our algorithms also perform well when the base station is inside the field. There have been no approaches so far for this scenario except direct transmission.

Although the simulations done in C makes us strongly believe that our algorithms will outperform others in a real environment, we will extend the network simulator ns to simulate PEDAP protocols in order to verify our results in a closer to real environment.

REFERENCES

- I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: A survey," *Computer Networks*, vol. 38, no. 4, pp. 393– 422, 2002.
- [2] A. Sinha and A. Chandrakasan, "Dynamic power management in wireless sensor networks," *Trans. Design and Test of Computers*, vol. 18, no. 2, pp. 62–74, Mar./Apr. 2001.
- [3] S. Singh, M. Woo, and C. S. Raghavendra, "Power-aware routing in mobile ad hoc networks," in *Mobile Computing and Networking*, 1998, pp. 181–190.
- [4] I. Stojmenovic and X. Lin., "Power-aware localized routing in wireless networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 12, no. 11, pp. 1122–1133, 2001.
- [5] J.-H. Chang and L. Tassiulas, "Energy conserving routing in wireless ad-hoc networks," in *IEEE INFOCOM 2000*, March 2000, pp. 22–31.
- [6] J. Chang and L. Tassiulas, "Maximum lifetime routing in wireless sensor networks," in *Proceedings of Advanced Telecommunications and Information Distribution Research Program, College Park, MD*, 2000.
- [7] M. Bhardwaj, A. Chandrakasan, and T. Garnett, "Upper bounds on the lifetime of sensor networks," in *IEEE International Conference on Communications*, 2001, pp. 785–790.
- [8] C. Intanagonwiwat, R. Govindan, and D. Estrin, "Directed diffusion: a scalable and robust communication paradigm for sensor networks," in Mobile Computing and Networking, 2000, pp. 56–67.
- [9] W. Heinzelman, J. Kulik, and H. Balakrishnan, "Adaptive protocols for information dissemination in wireless sensor networks," in *MOBICOM*, 1999, pp. 174–185.
- [10] K. Kalpakis, K. Dasgupta, and P. Namjoshi, "Maximum lifetime data gathering and aggregation in wireless sensor networks," in *Proceedings* of the 2002 IEEE International Conference on Networking (ICN'02), August 2002, pp. 685–696.
- [11] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in 33rd Annual Hawaii International Conference on System Sciences,, 2000, pp. 3005 3014.
- [12] S. Lindsey and C. S. Raghavendra, "Pegasis: Power-efficient gathering in sensor information systems," in *IEEE Aerospace Conference*, March 2002.
- [13] R. G. Gallager, P. A. Humblet, and P. M. Spira, "A distributed algorithm for minimum-weight spanning trees," ACM Transactions on Programming Languages and Systems, vol. 5, no. 1, pp. 66–77, January 1983.