# PEGASIS: Power-Efficient Gathering in Sensor Information Systems

**Stephanie Lindsey**
**Cauligi S. Raghavendra**

**Corresponding Author: Cauligi S. Raghavendra**
**Computer Systems Research Department**
**The Aerospace Corporation**
**P.O. Box 92957**
**Los Angeles, CA 90009-2957**
**raghu@aero.org**
**work: 310-336-1686**
**fax: 310-336-4402**

**Key Words: Sensor Networks, Wireless Communications, Data Collection, Protocols**

# PEGASIS: Power-Efficient GAthering in Sensor Information Systems

**Stephanie Lindsey and Cauligi S. Raghavendra**

**Computer Systems Research Department**
**The Aerospace Corporation**
**P. O. Box 92957**
**Los Angeles, CA 90009-2957**

## Abstract

**Sensor webs consisting of nodes with limited battery power and wireless communications are deployed to collect useful information from the field. Gathering sensed information in an energy efficient manner is critical to operate the sensor network for a long period of time. In [3] a data collection problem is defined where, in a round of communication, each sensor node has a packet to be sent to the distant base station. If each node transmits its sensed data directly to the base station then it will deplete its power quickly. The LEACH protocol presented in [3] is an elegant solution where clusters are formed to fuse data before transmitting to the base station. By randomizing the cluster heads chosen to transmit to the base station, LEACH achieves a factor of 8 improvement compared to direct transmissions, as measured in terms of when nodes die. In this paper, we propose PEGASIS (Power-Efficient GAthering in Sensor Information Systems), a near optimal chain-based protocol that is an improvement over LEACH. In PEGASIS, each node communicates only with a close neighbor and takes turns transmitting to the base station, thus reducing the amount of energy spent per round. Simulation results show that PEGASIS performs better than LEACH by about 100 to 300% when 1%, 20%, 50% , and 100% of nodes die for different network sizes and topologies.**

## 1. Introduction

Inexpensive sensors capable of significant computation and wireless communications are becoming available [2,4]. A web of sensor nodes can be deployed to collect useful information from the field, for example, in harsh physical environments [13]. These sensor nodes collect audio, seismic, and other types of data and collaborate to perform a high level task in the network. Sensor nodes are severely constrained by the amount of battery power available, limiting the lifetime and quality of the network. Since wireless communications consume significant amounts of battery power, sensor nodes should spend as little energy as possible receiving and transmitting data [5,10,12]. It is necessary for communication protocols to maximize nodes' lifetimes [9], reduce bandwidth consumption by using local collaboration among the nodes, and tolerate node failures [14].

Figure 1 shows a 100-node sensor network in a play field of size 50m x 50m. A typical application in a sensor web is gathering of sensed data at a distant base station (BS) [3]. Each sensor node has power control and the ability to transmit data to any other sensor node or directly to the BS [6,7]. We assume that all nodes have location information about all other nodes. However, if this were not the case, our scheme would still work. Nodes would have to expend some extra energy to find their close neighbors. They could do this by sending with enough power to signal a node, and then gradually reduce its power to find which neighbor is closest to it. In this paper, our model sensor network has the following properties:

- The BS is fixed at a far distance from the sensor nodes.
- The sensor nodes are homogeneous and energy constrained with uniform energy.
- No mobility of sensor nodes.

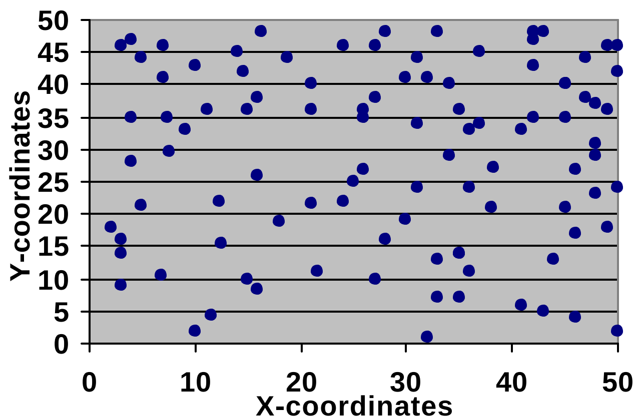


**Figure 1.** Random 100-node topology for a 50m x 50m network. BS is located at (25, 150), which is at least 100m from the nearest node.

In each round of this data-gathering application, all data from all nodes need to be collected and transmitted to the BS, where the end-user can access the data. A simple approach to accomplish this task is for each node to transmit its data directly to the BS. Since the BS is located far away, the cost to transmit to the BS from any node is high and nodes will die very quickly. Therefore, an improved approach is to use as few transmissions as possible to the BS and minimize the amount of data that must be transmitted to the BS.

In sensor networks, data fusion helps to reduce the amount of data transmitted between sensor nodes and the BS. Data fusion combines one or more data packets from different sensor measurements to produce a single packet as described in [3]. The LEACH protocol presented in [3] is an elegant solution to this data collection problem, where a small number of clusters are formed in a self-organized manner. A designated node in each cluster collects and fuses data from nodes in its cluster and transmits the result to the BS. LEACH uses randomization to rotate the cluster heads and achieves a factor of 8 improvement compared to the direct approach, before the first node dies. Further improvements can be obtained if each node communicates only with close neighbors, and only one designated node sends the combined data to the BS in each round.

In this paper we present an improved protocol called PEGASIS (Power-Efficient GAthering in Sensor Information Systems), which is near optimal for this data gathering application in sensor networks. The key idea in PEGASIS is to form a chain among the sensor nodes so that each node will receive from and transmit to a close neighbor. Gathered data moves from node to node, get fused, and eventually a designated node transmits to the BS. Nodes take turns transmitting to the BS so that the average energy spent by each node per round is reduced. Building a chain to minimize the total length is similar to the traveling salesman problem, which is known to be intractable. However, with the radio communication energy parameters, a simple chain built with a greedy approach performs quite well. The PEGASIS protocol achieves between 100 to 300% improvement when 1%, 20%, 50% and 100% of nodes node die compared to the LEACH protocol.

Our scheme can be modified appropriately if some of the stated assumptions about sensor nodes are not valid. If nodes are not within transmission range of each other, then alternative, possibly multi-hop transmission paths will have to be used. In fact, our chain based schemes will not be affected that much as each node communicates only with a local neighbor and we can use a multi-hop path to transmit to the BS. We need to make some adjustments in the chain construction procedure to ensure that no node is left out. The s LEACH protocol relies on direct reachability to function correctly. To ensure balanced energy dissipation in the network, an additional parameter could be considered to compensate for nodes that must do more work every round. If the sensor nodes have different initial energy levels, then we could consider the remaining energy level for each node in addition to the energy cost of the transmissions. The assumption of location information is not critical. The BS can determine the locations and transmit to all nodes, or the nodes can determine this through received signal strengths. For example, nodes could transmit progressively reduced signal strengths to find a close neighbor to exchange data. This would require the nodes to consume some energy when trying to find local neighbors, however, this is only a fixed initial energy cost when constructing the chain. If nodes are mobile, then different methods of transmission could be examined. For instance, if nodes could approximate how often and at what speed other nodes are moving, then it could determine more intelligently how much power is needed to reach the other nodes. Perhaps, the BS can help coordinate the activities of nodes in data transmissions. Discussion of schemes with mobile sensor nodes is beyond the scope of this paper.

## 2. Radio Model for PEGASIS

We use the same radio model as discussed in [3] which is the first order radio model. In this model, a radio dissipates $E_{elec}$ = 50 nJ/bit to run the transmitter or receiver circuitry and $\epsilon_{amp}$ = 100 pJ/bit/m$^2$ for the transmitter amplifier. The radios have power control and can expend the minimum required energy to reach the intended recipients. The radios can be turned off to avoid receiving unintended transmissions.

An $r^2$ energy loss is used due to channel transmission [8,11]. The equations used to calculate transmission costs and receiving costs for a *k*-bit message and a distance *d* are shown below:

**Transmitting**

$$E_{Tx}(k, d) = E_{Tx-elec}(k) + E_{Tx-amp}(k,d)$$

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

**Receiving**

$$E_{Rx}(k) = E_{Rx-elec}(k)$$

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

Receiving is also a high cost operation, therefore, the number of receives and transmissions should be minimal.

LEACH and PEGASIS use the same constants ($E_{elec}$, $\epsilon_{amp}$, and k) for calculating energy costs, therefore the PEGASIS achieves its energy savings by minimizing d and the number of transmissions and receives for each node. Therefore, for a $d^4$ model, PEGASIS would achieve even greater savings compared to LEACH.

In our simulations, we used a packet length *k* of 2000 bits. With these radio parameters, when $d^2$ is 500, the energy spent in the amplifier part equals the energy spent in the electronics part, and therefore, the cost to transmit a packet will be twice the cost to receive.

It is assumed that the radio channel is symmetric so that the energy required to transmit a message from node *i* to node *j* is the same as energy required to transmit a message from node *j* to node *i* for a given signal to noise ratio (SNR).

## 3. Energy Cost Analysis for Data Gathering

In this section we will analyze the cost of data gathering from a sensor web to the distant BS. Recall that the data collection problem of interest is to send a *k*-bit packet from each sensor node in each round. Of course, the goal is to keep the sensor web operating as long as possible. A fixed amount of energy is spent in receiving and transmitting a packet in the electronics, and an additional amount proportional to $d^2$ is spent while transmitting a packet. There is also a cost of 5 nJ/bit/message for data fusion.

With the direct approach, all nodes transmit directly to the BS which is usually located very far away. Therefore, every node will consume a significant amount of power to transmit to the BS in each round. Since the nodes have a limited amount of energy, nodes will die quickly, causing the reduction of the system lifetime. As observed in [3], the direct approach would work best if the BS is located close to the sensor nodes or the cost of receiving is very high compared to the cost of transmitting data.

For the rest of the analysis, we assume a 100-node sensor network with the BS located far away. In this scenario, energy costs can be reduced if the data is gathered locally among the sensor nodes and only a few nodes transmit the fused data to the BS. This is the approach taken in LEACH, where clusters are formed dynamically in each round and cluster-heads (leaders for each cluster) gather data locally and then transmit to the BS. Cluster-heads are chosen randomly, but all nodes have a chance to become a cluster-head in LEACH, to balance the energy spent per round by each sensor node. For a 100-node network in a 50m x 50m field with the BS located at (25,150), which is at least 100m from the closest node, LEACH achieves a factor of 8 improvement compared to the direct approach in terms of number of rounds before the first node dies.

Although this approach is about 8x better than the direct transmission, there is still some room to save even more energy. The cost of the overhead to form the clusters is expensive. In LEACH, in every round 5% of nodes are cluster-heads, and they must broadcast a signal to reach all nodes. In addition, several cluster-heads transmit the fused data from the cluster to the distant BS. Further improvement in energy cost for data gathering can be achieved if only one node transmits to the BS per round and if each node transmits only to local neighbors in the data fusion phase. This is done in the PEGASIS protocol to obtain an additional factor of 2 or more improvement compared to LEACH.

For the 100-node network shown in Figure 1, we can determine a bound on the maximum number of rounds possible before the first node dies. In each round, every node must transmit their packet and some node must receive it. So, each node spends two times the energy cost for electronics and some additional cost depending on how far a node transmits. Since some node must transmit the fused message to the BS in each round, on the average each node must incur this cost at least once every 100 rounds. With the energy cost parameters and the dimensions of play field in Figure 1, we can calculate the maximum rounds possible. The energy spent in each node for 100 rounds is about 100*.0002 Joules for the electronics and at least .002 Joules for one message transmission to the BS. With an initial energy in each node to be .25 Joules, the maximum number of rounds possible before a node dies is approximately 1100. The actual number will be less since we did not account for the energy spent in a node for local transmission, which depends on distance, and the cost for data fusion. Therefore, the upper bound will likely be less than 1000 rounds. The PEGASIS protocol achieves about 800 rounds, which is near optimal.

## 4. PEGASIS: Power-Efficient GAthering in Sensor Information Systems

The main idea in PEGASIS is for each node to receive from and transmit to close neighbors and take turns being the leader for transmission to the BS. This approach will distribute the energy load evenly among the sensor nodes in the network. We initially place the nodes randomly in the play field, and therefore, the *i* –th node is at a random location. The nodes will be organized to form a chain, which can either be accomplished by the sensor nodes themselves using a greedy algorithm starting from some node. Alternatively, the BS can compute this chain and broadcast it to all the sensor nodes.

We used random 100-node networks for our simulations with similar parameters used in [3]. We placed the BS at a far distance from all other nodes. For a 50m x 50m plot, our BS is located at (25, 150) so that the BS is at least 100m from the closest sensor node.

For constructing the chain, we assume that all nodes have global knowledge of the network and employ the greedy algorithm. We could have constructed a loop, however, to ensure that all nodes have close neighbors is difficult as this problem is similar to the traveling salesman problem. The greedy approach to constructing the chain works well and this is done before the first round of communication. To construct the chain, we start with the furthest node from the BS. We begin with this node in order to make sure that nodes farther from the BS have close neighbors, as in the greedy algorithm the neighbor distances will increase gradually since nodes already on the chain cannot be revisited. Figure 2 shows node 0 connecting to node 3, node 3 connecting to node 1, and node 1 connecting to node 2 in that order. When a node dies, the chain is reconstructed in the same manner to bypass the dead node.

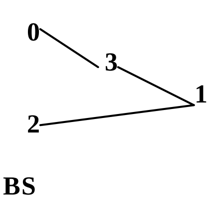


**Figure 2.** Chain construction using the greedy algorithm.

For gathering data in each round, each node receives data from one neighbor, fuses with its own data, and transmits to the other neighbor on the chain. Note that node *i* will be in some random position *j* on the chain. Nodes take turns transmitting to the BS, and we will use node number *i* mod N (N represents the number of nodes) to transmit to the BS in round *i*. Thus, the leader in each round of communication will be at a random position on the chain, which is important for nodes to die at random locations. The idea in nodes dieing at random places is to make the sensor network robust to failures. In a given round, we can use a simple control token passing approach initiated by the leader to start the data transmission from the ends of the chain. The cost is very small since the token size is very small. In Figure 3, node c2 is the leader, and it will pass the token along the chain to node c0. Node c0 will pass its data towards node c2. After node c2 receives data from node c1, it will pass the token to node c4, and node c4 will pass its data towards node c2.

c0→c1→c2←c3←c4
↓
BS

**Figure 3.** Token passing approach.

PEGASIS performs data fusion at every node except the end nodes in the chain. Each node will fuse its neighbor's data with its own to generate a single packet of the same length and then transmit that to its other neighbor (if it has two neighbors). In the above example, node c0 will pass its data to node c1. Node c1 fuses node c0's data with its own and then transmits to the leader. After node c2 passes the token to node c4, node c4 transmits its data to node c3. Node c3 fuses node c4's data with its own and then transmits to the leader. Node c2 waits to receive data from both neighbors and then fuses its data with its neighbors' data. Finally, node c2 transmits one message to the BS.

Thus, in PEGASIS each node will receive and transmit one packet in each round and be the leader once every 100 rounds. With our simulation experiments, we found that the greedy chain construction performs well with different size networks and random node placements. In constructing the chain, it is possible that some nodes may have relatively distant neighbors along the chain. Such nodes will dissipate more energy in each round compared to other sensors. We improved the performance of PEGASIS by not allowing such nodes to become leaders. We accomplished this by setting a threshold on neighbor distance to be leaders. Table 1 reflects this improvement. We may be able to slightly improve PEGASIS's performance further by applying a threshold adaptive to the remaining energy levels in nodes. Whenever a node dies, the chain will be reconstructed and the threshold can be changed to determine which nodes can be leaders.

PEGASIS improves on LEACH by saving energy in several stages. First, in the local gathering, the distances that most of the nodes transmit are much less compared to transmitting to a cluster-head in LEACH. Second, the amount of data for the leader to receive is at most two messages instead of 20 (20 nodes per cluster in LEACH for a 100-node network). Finally, only one node transmits to the BS in each round of communication.

## 5. Experimental Results

To evaluate the performance of PEGASIS, we simulated PEGASIS and LEACH using several random 100-node networks. Figure 1 shows a random 100-node network. The BS is located at (25, 150) in a 50m x 50m field, and the BS is located at (50,300) in a 100m x 100m field. We ran the simulations to determine the number of rounds of communication when 1%, 20%, 50% and 100% of the nodes die using direct transmission, LEACH, and PEGASIS with each node having the same initial energy level. Once a node dies it is considered dead for the rest of the simulation. Our simulations show that PEGASIS achieves:

- approximately 2x the number of rounds compared to LEACH when 1%, 20%, 50%, and 100% of nodes die for a 50m x 50m network.
- approximately 3x the number of rounds compared to LEACH when 1%, 20%, 50%, and 100% nodes die for a 100m x 100m network.
- balanced energy dissipation among the sensor nodes to have full use of the complete sensor network.
- near optimal performance.

Table 1 summarizes the results with initial energy per node of 0.25J, .5J, and 1.0J for the 50m x 50m and 100m x 100m networks. The shaded portion is for the 50m x 50m network. The nodes begin to die at a more uniform rate after about 20% nodes die. This is because the distances between the nodes become greater, and nodes have to become leaders more often causing the energy to drain rapidly. As can be expected, the number of rounds doubles as the energy/node doubles for a given size of the network.

**Table 1.** Number of rounds when 1%, 20%, 50%, and 100% nodes die. The shaded portion represents a 50m x 50m network, and the non-shaded portion represents a 100m x 100m network.

| Energy J/Node | Protocol | 1% | 20% | 50% | 100% |
|---|---|---|---|---|---|
| | Direct | 54 | 62 | 76 | 117 |
| .25 | LEACH | 402 | 480 | 523 | 635 |
| | PEGASIS | 788 | 1004 | 1041 | 1096 |
| | Direct | 108 | 124 | 152 | 235 |
| .5 | LEACH | 803 | 962 | 1036 | 1208 |
| | PEGASIS | 1578 | 2011 | 2082 | 2192 |
| | Direct | 215 | 248 | 304 | 471 |
| 1.0 | LEACH | 1610 | 1921 | 2055 | 2351 |
| | PEGASIS | 3159 | 4023 | 4165 | 4379 |
| | Direct | 14 | 16 | 20 | 30 |
| .25 | LEACH | 166 | 204 | 232 | 308 |
| | PEGASIS | 335 | 624 | 684 | 779 |
| | Direct | 28 | 32 | 40 | 61 |
| .5 | LEACH | 339 | 408 | 461 | 576 |
| | PEGASIS | 675 | 1250 | 1362 | 1544 |
| | Direct | 56 | 64 | 80 | 122 |
| 1.0 | LEACH | 690 | 812 | 911 | 1077 |
| | PEGASIS | 1346 | 2497 | 2720 | 3076 |

Figure 4 shows the number or rounds until 1%, 20%, 50%, 100% nodes die for a 50m x 50m network and Figure 5 shows same parameters but for a 100m x 100m network. PEGASIS is approximately 2x better than LEACH in all cases for a 50m x 50m network. The initial energy value for nodes is 0.25J in Figure 4 and 0.50J in Figure 5. As the energy level doubles the number of rounds also doubles for all cases. For a 100m x 100m network, PEGASIS performs about 3x better than LEACH.

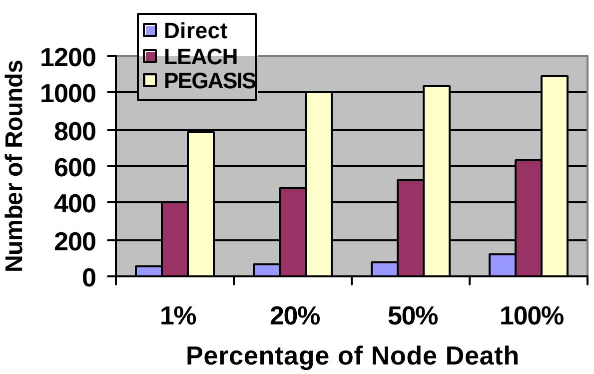


**Figure 4.** Performance results for a 50m x 50m network with initial energy .25J/node.

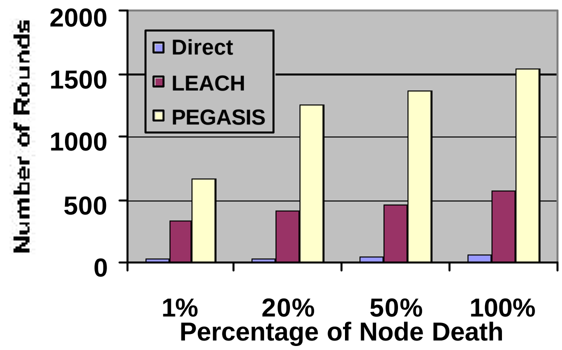


**Figure 5.** Performance results for a 100m x 100m network with initial energy .5J/node.

## 6. Conclusions and Future Work

In this paper, we describe PEGASIS, a greedy chain protocol that is near optimal for a data-gathering problem in sensor networks. PEGASIS outperforms LEACH by eliminating the overhead of dynamic cluster formation, minimizing the distance non leader-nodes must transmit, limiting the number of transmissions and receives among all nodes, and using only one transmission to the BS per round. Nodes take turns to transmit the fused data to the BS to balance the energy depletion in the network and preserves robustness of the sensor web as nodes die at random locations.

Distributing the energy load among the nodes increases the lifetime and quality of the network. Our simulations show that PEGASIS performs better than LEACH by about 100 to 300% when 1%, 20%, 50%, and 100% of nodes die for different network sizes and topologies. PEGASIS shows an even further improvement as the size of the network increases.

In order to verify our assumptions about PEGASIS, we will extend the network simulator ns-2 to simulate PEGASIS, LEACH, and direct transmission protocols. Based on our C simulations, we expect that PEGASIS will outperform the other two protocols in terms of system lifetime and the quality of the network.

## References

[1] Bluetooth Initiative. http://www.bluetooth.com, 1999.

[2] D. Estrin, R. Govindan, J. Heidemann, and Satish Kumar. Next Century Challenges: Scalable Coordination in Sensor Networks. In *Proceedings of Mobicom' 99*, 1999.

[3] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan. Energy-Efficient Communication Protocol for Wireless Microsensor Networks. In *Proceedings of the Hawaii Conference on System Sciences*, Jan. 2000.

[4] J. Kulik, W. Rabiner, and H. Balakrishnan. Adaptive Protocols for Information Dissemination in Wireless Sensor Networks. In *Proceedings of Mobicom' 99*, 1999.

[5] W. Mangione-Smith and P.S. Ghang. A Low Power Medium Access Control Protocol for Portable Multi-Media Systems. In *Proceedings 3rd Intl. Workshop on Mobile Multimedia Communications*. Princeton, NJ, Sept. 25-27, 1996.

[6] R. Pichna and Q. Wang. Power Control. In *The Mobile Communications Handbook*. CRC Press, 1996, pp. 370-380.

[7] R. Ramanathan and R. Hain. Topology Control of Multihop Wireless Networks Using Transmit Power Adjustment. In *Proceedings Infocom 2000*, 2000.

[8] T. S. Rappaport. *Wireless Communications*. Prentice-Hall, 1996.

[9] S. Singh, M. Woo and C.S. Raghavendra. Power-Aware Routing in Mobile Ad Hoc Networks. *In Proceedings ACM/IEEE Mobicom'98*, 1998.

[10] K. M. Sivalingam, M. B. Srivastava and P. Agrawal. Low Power Link and Access Protocols for Wireless Multimedia Networks. In *Proceedings IEEE Vehicular Technology Conference VTC'97*, May 1997.

[11] R. Steele. *Mobile Radio Communications*. Pentech Press, London, 1992.

[12] M. Stemm, P. Gauthier, D. Harada and R. Katz. Reducing Power Consumption of Network Interfaces in Hand-Held Devices. In *Proceedings 3rd Intl. Workshop on Mobile Multimedia Communications*, Sept. 25-27, 1996, Princeton, NJ.

[13] The WINS Project. http://www.janet.ucla.edu/WINS.

[14] M. Zorzi and R. R. Rao. Energy Management in Wireless Communications. In *Proceedings 6th WINLAB Workshop on Third Generation Wireless Information Networks*, March 1997.