

# Energy Efficient Algorithm: PEDAP for WSN

Sanjoy Sana  
M.Tech, Department of IT, Hindustan University  
Chennai, India

Dr.K.Ramesh Kumar  
Associate professor (IT), Hindustan University  
Chennai, India

**Abstract**— Wireless sensor network (WSNs) are networks of distributed autonomous device that can sense or monitor physical or environmental condition cooperatively. The WNS application can be classified in different categories: monitoring, alerting, healthcare assistance, and actuating, among others. In this 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. PEDAP achieves between 4x to 20x improvement in network lifetime compared with LEACH, and about three times improvement compared with PEGASIS.

**Keywords**- Energy Efficiency, Data fusion and Aggregation, PEGASIS, LEACH, PEDAP.

## I. INTRODUCTION

WIRELESS SENSOR NETWORKS (WSNs) have gained a great deal of attention in recent years, particularly with the development of smart sensors. Smart nodes are low-power devices equipped with one or more sensors, a processor, memory, a power supply, a radio chipset, and some kinds of actuators. In this sense, WSNs are oriented to enable the distributed sensing of a physical phenomenon to perform distributed sensing and actuation tasks. 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 [1]. 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.

The wireless sensor node, being a micro-electronic device, can only be equipped with a limited power source (<0.5 Ah, 1.2 V). In some application scenarios, replenishment of power resources might be impossible. Sensor node lifetime, therefore, shows a strong dependence on battery lifetime. In a multihop ad hoc sensor network, each node plays the dual role of data originator and data router. The disfunctioning of few nodes can cause significant topological changes and might require re-routing of packets and re-organization of the network. Hence, power conservation and power management take on additional importance. It is for these reasons that researchers are currently focusing on the design of power-aware protocols and algorithms for sensor networks.

The main task of a sensor node in a sensor field is to detect events, perform quick local data processing, and then transmit the data. Power consumption can hence be divided into three domains: sensing, communication, and data processing. Sensing power varies with the nature of applications. Sporadic sensing might consume lesser power than constant event monitoring. The complexity of event detection also plays a crucial role in determining energy expenditure. Higher ambient noise levels might cause significant corruption and increase detection Complexity.

A sensor node expends maximum energy in data communication. This involves both data transmission and reception. It can be shown that for short-range communication with low radiation power ( $-10$  dbm), transmission and reception energy costs are nearly the same. Mixers, frequency synthesizers, voltage control oscillators, phase locked loops (PLL) and power amplifiers, all consume valuable power in the transceiver circuitry. It is important that in this computation we not only consider the active power but also the start-up power consumption in the transceiver circuitry. The start-up time, being of the order of hundreds of micro-seconds, makes the start-up power non-negligible. This high value for the start-up time can be attributed to the lock time of the PLL. As the transmission packet size is reduced, the start-up power consumption starts to dominate the active power consumption[3]. As a result, it is inefficient in turning the transceiver ON and OFF, because a large amount of power is spent in turning the transceiver back ON each time. In present a formulation for the radio power consumption ( $P_c$ ) as

$$P_c = N_T [P_T(T_{on} + T_{st}) + P_{out}(T_{on})] + N_R [P_R(R_{on} + R_{st})]$$

where  $P_{T/R}$  is the power consumed by the transmitter/ receiver;  $P_{out}$ , the output power of the transmitter;  $T/R_{on}$ , the transmitter/receiver on time;  $T/R_{st}$ , the transmitter/receiver start-up time and  $N_{T/R}$ , the number of times transmitter/receiver is switched on per unit time, which depends on the task and medium access control (MAC) scheme used.  $T_{on}$  can further be rewritten as  $L/R$ , where  $L$  is the packet size and  $R$ , the data rate. Today's state-of-the-art low power radio transceiver has typical  $P_T$  and  $P_R$  values around 20 dbm and  $P_{out}$  close to 0 dbm [59]. Note that Pico Radio aims at a  $P_c$  value of  $\sim 20$  dbm.

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  $nk$ .

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 and PEGASIS are proposed. In LEACH, the key idea is to reduce the number of nodes communicating directly with the base station[2]. The protocol achieves this by forming a small number of clusters in a self-organizing 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 [4].

The rest of the paper is organized as follows. In Section II, we formulate our system model and the data gathering problem. The PEDAP protocols are described in detail in Section III. Next, the feasibility of implementation of our algorithms is discussed in Section IV. Finally, we conclude the paper and present future research directions in Section V.

## 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. In the specified radio model, the energy dissipation of the radio in order to run the transmitter or receiver circuitry is equal to  $E_{elec} = 50n Jbit$ , and to run the transmit amplifier it is equal to  $E_{amp} = 100p Jbitm^2$ . It is also assumed an  $r2$  energy loss due to channel transmission[5]. Therefore, the energy expended to transmit a  $k$ -bit packet to a distance  $d$  and to receive that packet with this radio model is:

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

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

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, 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.

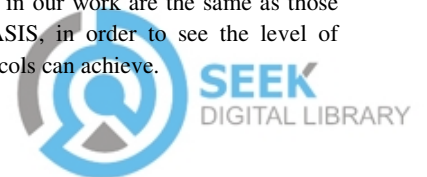
It is also important to note that the cost of one transmission of a  $k$ -bit packet to the system is either:

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

$$C'_i(k) = E_{elec} * k + E_{amp} * k * d_{ib}^2 \quad (4)$$

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

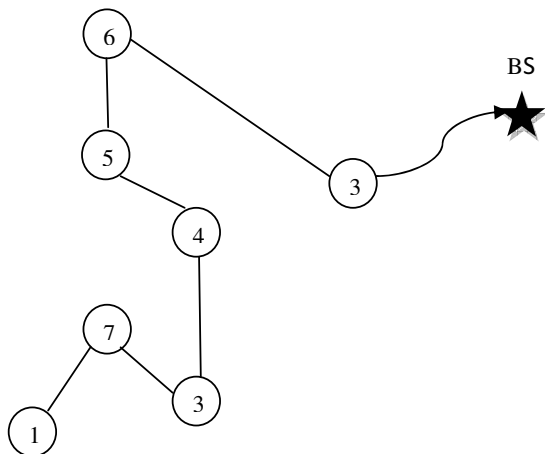


Fig. 1. Chain based routing scheme on a sample network.

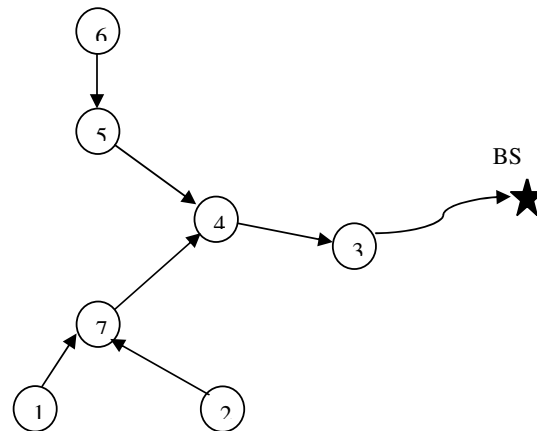


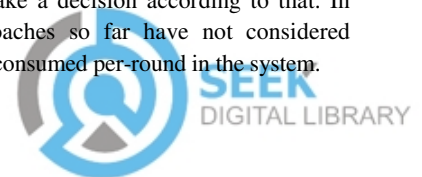
Fig. 2. Minimum spanning tree based routing scheme on a sample network.

**C. Energy Analysis for Data Routing**

In 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 (cluster-heads) 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 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[6]. 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 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[10]. The first part of the idea can be realized optimally by computing a minimum spanning tree over the sensor 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[7]. 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:

$$C_{ij}(k) = 2 \frac{E_{elec} * k + E_{amp} * k * d_{ij}^2}{e_i} \quad (5)$$

$$C'_i(k) = \frac{E_{elec} * k + E_{amp} * k * d_{ib}^2}{e_i} \quad (6)$$

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.  $0 \leq e_i \leq 1$ ).

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 PEDAPPA (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 consuming system[9]. Besides knowing the locations of the nodes, the base station can also estimate the remaining energy levels of the nodes by using the given cost model, since it knows how much energy a node spends in a round. After some certain number of rounds passed (e.g. 100) the base station re-computes the routing information excluding the dead nodes. After each computation, the base station sends each node the required information for that node (i.e. the node's parent in the tree in order to reach to the base station; the time slot number when the node will send its data to its parent in a round; from how many different neighbors the node will receive packets in a round and when; etc.). So, the cost of setting-up the system with the new routing information is equal to only the sum of costs of running the receiver circuitry of each node. Therefore, the set-up cost for periodically establishing the scheme is very small compared to LEACH and PEGASIS.



For the two algorithms proposed in this work, the protocols are the same [8]. 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. 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 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 assuming the visibility graph is given.

#### V. 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. 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.

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.

#### Reference

- [1]. M. C. Rodríguez-Sánchez, S. Borromeo, *Member, IEEE*, and J. A. Hernández-Tamames. "Wireless Sensor Networks for Conservation and Monitoring Cultural Assets," *IEEE Sensors Journal*, vol. 11, no. 6, June 2011.
- [2]. Marco F. Duarte, Godwin Shen, Antonio Ortega, and Richard G. Baraniuk, "Signal Compression in Wireless Sensor Networks".
- [3]. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: A survey," *Computer Networks*, vol. 38, no. 4, pp. 393–422, 2002.
- [4]. Jonathan Henderson. "A Study of the LEACH Protocol for Wireless Sensor Networks." Mississippi Valley State University, Itta Bena, Mississippi 38941.
- [5]. Wendi Rabiner Heinzelman, Anantha Chandrakasan, and Hari Balakrishnan. "Energy-Efficient Communication Protocol for Wireless Microsensor Networks," 33rd Hawaii International Conference on System Sciences – 2000.
- [6]. Hu'seyin O' zgu'r Tan and I'brahim Ko'rpeog'lu "Power Efficient Data Gathering and Aggregation in Wireless Sensor Networks", Department of Computer Engineering, Bilkent University.
- [7]. W. Heinzelman, A. Chandrakasan and H. Balakrishnan, "Energy-Efficient Communication Protocols for Wireless Microsensor Networks," *Proceedings of the Hawaiian International Conference on Systems Science*, January 2000.
- [8]. S. Lindsey and C. S. Raghavendra, "Pegasis: Power-efficient gathering in sensor information systems," in *IEEE Aerospace Conference*, March 2002.
- [9]. John A. Stankovic, Tarek F. Abdelzaher, Chenyng Lu, Lui Sha and Jennifer C. Hou, "Real-Time Communication and Coordination in Embedded Sensor Networks", *Proceedings of the IEEE*, V01.91. No.7. pp. 1002-1027. July. 2003.
- [10]. J. L. Hill and D. E. Culler, "Mica: A wireless platform for deeply embedded networks," *IEEE Micro*, vol. 22, no. 6, pp. 12–24, Dec. 2002.
- [11]. T. Meng and R. Volkan. Distributed Network Protocols for Wireless Communication. In *Proc. IEEE ISCAS*, May 1998.

