Optimal Cluster Head Selection in the LEACH Architecture

Haiming Yang
Department of ECSE
Rensselaer Polytechnic Institute
Troy, NY 12180 U.S.A
yangh4@rpi.edu

Biplab Sikdar
Department of ECSE
Rensselaer Polytechnic Institute
Troy, NY 12180 U.S.A
bsikdar@ecse.rpi.edu

Abstract

LEACH (Low Energy Adaptive Clustering Hierarchy) [1] is one of the popular cluster-based structures, which has been widely proposed in wireless sensor networks. LEACH uses a TDMA based MAC protocol, and in order to maintain a balanced energy consumption, suggests that each node probabilistically become a cluster head. To reduce the energy consumption and to avoid the strict synchronization requirements of TDMA, we first apply a sleep-wakeup based decentralized MAC protocol to LEACH, then we present an analytic framework for obtaining the optimal probability with which a node becomes a cluster head in order to minimize the network's energy consumption. The analysis is first presented for small networks, under the assumption of identical expected distance of all cluster heads from the sink. Then the analysis is extended for large networks to consider the case when the distances of various sections of the network from the sink may be different, since nodes further away have to spend greater energy in order to reach the sink. Our simulation results show that using this optimal probability results in much more efficient energy consumption and compared with the current LEACH, our proposal consumes significantly less power.

1 Introduction

Wireless Sensor Networks (WSNs) can offer unique benefits and versatility with respect to low-power and low-cost rapid deployment for many applications which do not need human supervision. The nodes in WSNs are usually battery-operated sensing devices with limited energy resources and replacing or replenishing the batteries is usually not an option. Thus energy efficiency is one of the most important issues and designing power-efficient protocols is critical for prolonging the lifetime.

Clustering network is an efficient and scalable way to organize WSNs [8, 1]. A cluster head is responsible for conveying any information gathered by the nodes in its cluster and may aggregate and compress the data before transmitting it to the sink. However, this added responsibility results in a higher rate of energy drain at the cluster heads. One of the most popular clustering mechanisms, LEACH, addresses this by probabilistically rotating the role of cluster head among all nodes. However, unless each node selects its probability of becoming a cluster head wisely, the performance of the network may be far from optimal. The main focus of this paper is to address this problem and develop a framework to select the optimal probability with which a node should become a cluster head from the perspective of minimizing the network's energy consumption.

The rest of the paper is organized as follows. Section 2 presents a brief description of related work, Section 3 describes the problems with the current LEACH algorithm, and our MAC protocol for LEACH architecture. In section 4 we analyze the energy consumption of sensors and develop the framework for obtaining the optimal probability of cluster head selection. Finally, simulation results are presented in Section 5 while Section 6 concludes the paper.

2 Related Work

In sensor networks, the nodes serving as cluster heads are over-loaded with the long range transmissions to the remote nodes or sink as well as due to the extra processing necessary for data process. One way to ensure balance energy consumption is to rotate the role of a cluster head randomly and periodically over all the nodes as proposed in LEACH [1]. The operation of LEACH is divided into periods and each period consists of a set-up phase and a steady-state phase. During the setup phase, nodes communicate with short messages and are organized into clusters with some nodes selected as cluster heads. After the set-up phase, each cluster head sets up TDMA schedules for all leaf nodes in its cluster. Leaf nodes send any data they generate to the cluster heads according to the TDMA schedule. The cluster head aggregates and compresses this data before passing it to the sink. A number of schemes to prolong the lifetime of sensor networks have been proposed in literature. A new chain-based protocol called PEGASIS is proposed in [2] where each node communicates only with a close neighbor instead of directly communicating with the sink, and takes turns transmitting to the sink. A hybrid protocol APTEEN [3], which allows for comprehensive information retrieval, and the nodes in such a network not only react to time-critical situations, but give an overall picture of the network at periodic intervals. However, appropriate cluster head election mechanisms are required to reduce the energy consumption and enhance the lifetime of the network. In [8], each node computes the quotient of its own energy level and the aggregate energy remaining in the network. With this value each node decides if it becomes cluster head for this round or not. Nodes with higher power are more likely to become cluster heads than nodes with lower power. The problem with this scheme is that each node has to estimate the remaining energy in the network which requires additional communication with the sink and other nodes. In [4], the authors extend LEACH's stochastic cluster heads selection algorithm by adding a deterministic component, and change the probability of a node becoming a cluster head by the remaining energy of sensor nodes. In LEACH-C [6], the cluster formation is done at the beginning of each round using a centralized algorithm by the sink. Although the energy cost for cluster formation is higher in LEACH-C, the overall performance may be better than LEACH due to improved cluster formation by the sink. SMAC [7] is a kind of energy efficient MAC protocols for WSNs, which is based on the CSMA/CA [5] mechanism. In SMAC each node follows a periodic sleep-wakeup schedule, synchronized with its neighboring nodes. Although SMAC is energy efficient, it is not fit for cluster based architectures.

In this paper, we propose an optimal cluster head selection algorithm to prolong the lifetime of WSNs based on the LEACH architecture. We also present a more energy efficient MAC protocol based on SMAC for using in the LEACH architecture, replacing TDMA.

3 Protocol Description

There are several problems in the current LEACH, which makes the LEACH energy inefficient. Since cluster heads spend more energy than leaf nodes, it is quite important to reselect cluster heads periodically. If this probability is set high, more nodes will become cluster heads and the rate of energy consumption becomes high; If this probability is low, the size of each cluster becomes larger and the average distance between leaf nodes and their cluster head increases, and the rate of energy consumption in the network also increases since the energy consumption is related to the square of distance from leaf nodes from the cluster head. The LEACH protocol specifies that nodes become cluster heads with a probability of 5% under normal circumstances, a value which results in suboptimal operation in most scenarios. Also, the TDMA mechanism which is used by LEACH for intra-cluster communication does not scale when the number of nodes increases.

This paper addresses these two problems above by applying a sleep-wakeup based, decentralized MAC protocol to LEACH. We also analyze LEACH's cluster head selection algorithm to find the optimal probability.

The operation of MAC is divided into rounds. A round is defined as the period from an instance of cluster head selection to the next cluster head reselection. Each round consists of several cycles and each cycle consists of a setup phase, an intra cluster communication phase and an inter cluster communication phase. In each setup phase, the nodes may either exchange messages to select a new cluster head or exchange network setup and maintenance messages if there is no need to reselect a new cluster head. The cluster head selection is repeated once several cycles. During the intra cluster communication phase, cluster heads exchange data with the leaf nodes in their cluster and compress these data into a single message. After the intra cluster communication phase, the cycle comes to the inter cluster communication phase and only the the cluster heads keep the radio on to transmit the compressed data to the sink. During this period, each cluster head accesses the channel in distributed manner by using CSMA/CA. The cluster heads stay awake at all times while the leaf nodes may sleep during the inter cluster communications to save energy. Since CSMA/CA is used for both the intra and inter cluster communication, the distributed nature of the protocol reduces the complexity and improves the scalability as the number of nodes increases.

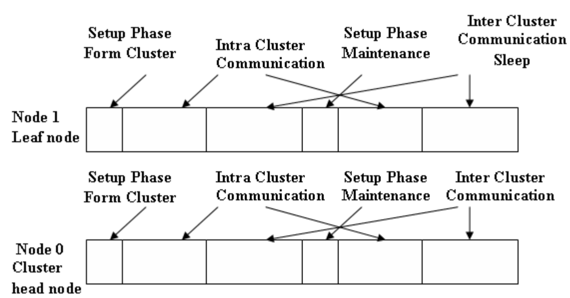

Figure 1. Protocol operation over two cycles

Figure 1 shows a simple example of the protocol's operation over two cycles for two nodes. In the first cycle, node 0 and node 1 both send cluster head election information in the setup phase, and node 0 is elected as the cluster head. After the setup phase, as a leaf node, node 1 sends data to node 0 in the intra cluster phase. Node 1 enters the sleep mode in the inter cluster phase when node 0 sends the data collected during the intra cluster phase to the sink. In the second cycle, there is no need to reselect the cluster head and thus in the setup phase, nodes in the cluster exchange network maintenance information. The setup phase is followed by the second intra cluster phase and inter cluster phase.

4 Optimal Cluster Head Selection

The calculations are based on the following assumptions and simplifications. We assume that the intra cluster communication phase is long enough, so all leaf nodes having data can send their data to the cluster head; And the inter cluster communication phase is long enough, so all head nodes having data can send their data to the sink. The cluster head performs data aggregation and compression before transmitting the data to the sink. The sink is assumed to be stationary and all sensor nodes are able to reach the sink. Finally, we assume symmetric propagation channels.

4.1 Radio Model for Energy Analysis

In this section, we describe our model for the energy consumed by the radios during transmission and reception. The nodes are assumed to have power control features so as to adjust their transmit power to the minimum level required for successful transmission. Now, different assumptions about the radio energy dissipation characteristics in the sending and receiving modes affect the performance of different protocols. To keep the model general, we assume that the radio dissipates $E_{elec}\ J/bit$ to run the transmitter or receiver circuit and $E_{amp}\ J/m^2$ for the transmitter amplifier to achieve an acceptable signal to noise ratio [1]. Assuming r^2 energy loss due to channel transmission, to send a k bits message to a distance of d meters using this radio model, the radio expends

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

Joules of energy and the radio expends

$$E_{Rx}(k, d) = kE_{elec} \tag{2}$$

Joules of energy to receive the message.

4.2 Energy Consumption Calculation

We consider a network area $L \times L$ meters with M sensor nodes randomly distributed with uniform distribution. The distance of the cluster heads to the sink is denoted by d and the distance between the leaf nodes in a cluster and their cluster head is denoted by r. For an arbitrary node in an arbitrary cluster, the probability that the node has data to send in an arbitrary cycle is denoted by p_{rate}, and the data sent by the nodes consists of k_{data} bytes. Since cluster heads compress the data before sending it to the sink, the length of this data is also assumed to be k_{data} bytes. We denote the expected number of nodes in a cluster by N with one cluster head and $N-1$ leaf nodes. Then the probability that the cluster head has data to send at the beginning of an inter cluster communication phase is given by $p_{hrate} = 1 - (1 - p_{rate})^N$. The time to transmit one byte on the network is denoted by T. The time devoted for intra cluster communications in each cycle is denoted by T_{intra} and the time devoted for inter cluster communication is denoted by T_{inter}. Since the setup phase has the same energy consumption characteristics as the intra cluster communication phase, we combine the setup phase and intra cluster communication phase together in T_{intra}. Also, the energy spent on receiving is almost the same as the energy spent on the idle listening, so the energy consumed during a period of idle listening can be calculated by evaluating the energy spent on receiving a message for the same amount of time.

The main objective is to minimize the mean energy consumption of every sensor node in each round. We denote the probability that a node serves as a cluster head in an arbitrary cycle by p and the energy consumption of a sensor node in a cycle when it is a cluster head by E_{head}. Thus the probability that a node is a leaf node in an arbitrary cycle is $1 - p$ and we denote the energy consumed by a leaf node in a cycle by E_{leaf}. Then the mean energy consumed by a node in a cycle is given by

$$E_{avg} = pE_{head} + (1 - p)E_{leaf} \tag{3}$$

A leaf node needs to turn on the radio only in the period of intra cluster communication. Then based on the radio model in Eqns. (1) and (2), the mean energy consumption of leaf node is given by

$$\begin{aligned} E_{leaf} &= p_{rate}\left[E_{Tx}(k_{data}, r) + E_{Rx}\left(\frac{T_{intra}}{T} - k_{data}\right)\right] \\ &\quad + (1 - p_{rate})\left[E_{Rx}\left(\frac{T_{intra}}{T}\right)\right] \end{aligned} \tag{4}$$

There are two parts of energy consumption in Eqn. (4). First, when the leaf node has data to send, an event which occurs with probability p_{rate}, the node spends $E_{Tx}(k_{data}, r)$ Joules to transmit the data to the cluster head. For the remaining period of the intra cluster communication phase, whose duration *in terms to bytes* that may be transmitted in this period is given by $\frac{T_{intra}}{T} - k_{data}$, the leaf node is in the idle listening mode. Since the energy consumed during idle listening is almost the same as the energy consumed during data receiving, the node consumes

$E_{Rx}(\frac{T_{intra}}{T} - k_{data})$ Joules. This is the first term of Eqn. (4). Second, in case the leaf node does not have any data to send to cluster head, which happens with probability $1 - p_{rate}$, the leaf node stays in the idle listening for the entire intra cluster communication period (whose duration in terms of bytes is given by $\frac{T_{intra}}{T}$). The energy spent in this case is given by $E_{Rx}(\frac{T_{intra}}{T})$ as shown in the second term of Eqn. (4). Note that since leaf nodes turn off their radios in the inter cluster period, there no energy consumed during this phase.

Cluster heads have to turn on their radio during both the intra cluster and the inter cluster periods. Using the radio model in Eqns. (1) and (2), the average energy consumed by a cluster head is given by

$$E_{head} = p_{hrate}\left[E_{Tx}(k_{data}, d) + E_{Rx}\left(\frac{T_{inter}}{T} - k_{data}\right)\right] +(1 - p_{hrate})\left[E_{Rx}\left(\frac{T_{inter}}{T}\right)\right] + E_{Rx}\left(\frac{T_{intra}}{T}\right) \quad (5)$$

With probability p_{hrate}, cluster heads have data to send to the sink during an inter cluster communication phase and this consumes $E_{Tx}(k_{data}, d)$ Joules energy. In the rest of the inter cluster communication period, of duration $\frac{T_{inter}}{T} - k_{data}$ in terms of bytes, the energy consumed is given by $E_{Rx}(\frac{T_{inter}}{T} - k_{data})$ since the cluster head is in the idle listen mode in this period. This is the first term in the equation above. With probability $1 - p_{hrate}$ the cluster head does not transmit any data to the sink and in this case the cluster head spends the entire inter cluster communication period in the idle listen mode, consuming $E_{Rx}(\frac{T_{inter}}{T})$ Joules and this is the second term in the equation above. Finally, the cluster head remains in the receiving mode during the intra cluster communication phase consuming $E_{Rx}(\frac{T_{intra}}{T})$ joules and this is the third term in Eqn. (5).

To evaluate Eqns. (4) and (5) we need to obtain the mean distance r between the leaf nodes in a cluster and their cluster head. To obtain r we need to characterize the area covered by a cluster and here we use the concept of Voronoi Tessellations [9]. Given a set of points in a plane, a Voronoi tessellation divides the plane into a set of polygonal regions, the boundaries of which are the perpendicular bisectors of the lines joining the points. Consider an open set $\Omega \in R^2$. The set $\{V_i\}_{i=1}^k$ is called a tessellation of Ω if $V_i \cap V_j = \varnothing$ for $i \neq j$ and $\cup_{i=1}^k V_j = \Omega$. Let d denote a distance defined on R^2. Given points $\{z_i\}_{i=1}^k$ belonging to Ω, the Voronoi region $\hat{V}_i$ corresponding to the point z_i is defined by

$$\hat{V}_i = \{x \in \Omega | d(x, z_i) < d(x, z_j) \text{ for } j = 1, \cdots, k, j \neq i\} \quad (6)$$

The points $\{z_i\}_{i=1}^k$ are the generators and the set $\{\hat{V}_i\}_{i=1}^k$ is a Voronoi tessellation of Ω. Each $\hat{V}_i$ is referred to as the Voronoi region corresponding to z_i. In the uniform, randomly deployed scenario considered in this paper, the cluster heads are the points z_i and the leaf nodes are points which fall into the Voronoi region $\hat{V}_i$ corresponding to the cluster head z_i. This is because leaf nodes will tend to get associated in the cluster belonging to the cluster head nearest to them. Since the nodes are uniformly distributed in the field, the mean position of cluster heads (x^*, y^*) is evenly located in the field, and the Voronoi region corresponding to each cluster head using standard Euclidean distance is a square region as shown in Figure 2.

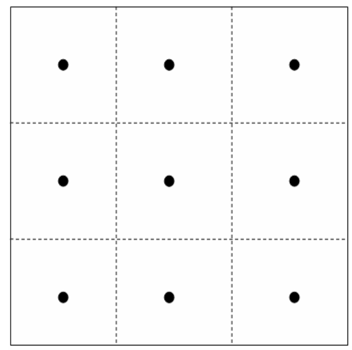

Figure 2. Voronoi tessellation separation

Since each node becomes a cluster head with probability p and the number of nodes in the network is M, the expected number of cluster heads in the network is Mp. Thus the expected area covered by each cluster is $\frac{L^2}{Mp}$ and as the regions in the Voronoi tessellation shown in Figure 2, this may be considered to be a $\sqrt{\frac{L^2}{Mp}} \times \sqrt{\frac{L^2}{Mp}}$ square area. Also, since the location (x, y) of a leaf node in a cluster is also uniformly and independently distributed in the cluster area, as is the position of the cluster head inside the cluster, (x^*, y^*). Thus we have $E[x] = E[x^*] = E[y] = E[y] = \frac{1}{2}\sqrt{\frac{L^2}{Mp}}$, and $E[x^2] = E[x^{*2}] = E[y^2] = E[y*^2] = \frac{1}{3}\frac{L^2}{Mp}$ and the average squared distance of a leaf node from its cluster head is given by

$$r^2 = E[(x - x^*)^2 + (y - y^*)^2] = \frac{L^2}{3Mp} \quad (7)$$

Now using the notation $k_{intra} = \frac{T_{intra}}{T}$ and $k_{inter} = \frac{T_{inter}}{T}$ and substituting Eqns. (4), (5) and (7) into Eqn. (3), the average energy consumed by a node in a cycle, E_{avg}, can be written as

$$E_{avg} = E_{amp} k_{data}\left[\frac{L^2(1-p)p_{rate}}{3Mp} + p(1-(1-p_{rate})^N)d^2\right] +E_{elec}(k_{intra} + p k_{inter}) \quad (8)$$

4.3 Analysis for Small Networks

For small networks or in a coarse grained analysis, the expected distance of each cluster head from the sink may be considered to be the same. In this case, every sensor node is assigned the same probability with which it becomes a cluster head. To obtain the optimal value of this probability, in Eqn. (8) we need to substitute the expected distance of a cluster head from the sink, d. The sink position (x_{sink}, y_{sink}) is fixed for a given sensor network scenario and the position of an arbitrary cluster head, (x^*, y^*), is uniformly distributed in the network area of $L \times L$ meters. Then, $E[x^*] = E[y^*] = \frac{L}{2}$ and $E[x^{*2}] = E[y^{*2}] = \frac{L^2}{3}$ and the average square of the distance between an arbitrary cluster head and the sink is

$$\begin{aligned} d^2 &= E[(x_{sink} - x^*)^2 + (y_{sink} - y^*)^2] \\ &= (L - \frac{1}{2}x_{sink})^2 + (L - \frac{1}{2}y_{sink})^2 - \frac{4}{3}L^2 \quad (9) \end{aligned}$$

Assuming the sink is located at $(0, 0)$, the expression for d further simplifies to

$$d^2 = L^2 + L^2 - \frac{4}{3}L^2 = \frac{2}{3}L^2 \quad (10)$$

Also, since every sensor node has the same probability p to become a cluster head, the expected number of nodes in a cluster is given by

$$N = \frac{1}{p} \quad (11)$$

Substituting this expression for N and the expression for d^2 from Eqn. (10) into Eqn (8), the average energy consumed by a sensor node in this case is given by

$$\begin{aligned} E_{avg} = & E_{amp}k_{data}L^2 \left[\frac{2p(1-(1-p_{rate})^{\frac{1}{p}})}{3} + \frac{(1-p)p_{rate}}{3Mp} \right] \\ & + E_{elec}[k_{intra} + pk_{inter}] \quad (12) \end{aligned}$$

In order to find the optimal value of the probability with which a node should become a cluster head that minimizes the mean energy consumption of the sensor nodes, we take the derivative of Eqn. (12) above with respect to p and equate it to zero. This is given by

$$\begin{aligned} 0 = & E_{elec}k_{inter} - \frac{p_{rate}E_{amp}k_{data}L^2}{3Mp^2} + \frac{2}{3}E_{amp}k_{data}L^2 \\ & + \frac{2E_{amp}k_{data}L^2(1-p_{rate})^{\frac{1}{p}}}{3} \left[\frac{\ln(1-p_{rate})}{p} - 1 \right] (13) \end{aligned}$$

Eqn. (13) can be easily resolved numerically, to obtain the optimal probability for a node to become a cluster head for a given rate of data generation, p_{rate}. We note that the optimal value of p also depends on the dimensions of the area covered by the network, the number of sensor nodes in the network, the length of messages, the rate at which data is generated at each node and the length of time a leaf node may sleep in a cycle, i.e., the length of the inter cluster communication period. For a given sensor hardware, E_{elec} and E_{amp} are constant values. Also, the dimension of sensor field, the number of nodes in the network and k_{data} and k_{inter} are also fixed when the application is decided. The optimal probability of sensor nodes to become cluster head can thus be easily derived from Eqn. (13) as a function of p_{rate}.

We now obtain an explicit solution for Eqn. (13) for the special case where all nodes have data to send in each cycle ($p_{rate} = 1$). In this case, the solution of Eqn. (13) for p yields

$$p = \sqrt{\frac{E_{amp}k_{data}L^2}{M(3E_{elec}k_{inter} + 2E_{amp}L^2k_{data})}} \quad (14)$$

as the optimal value of p.

4.4 Analysis for Large Networks

We now consider the case when the network is large and thus the distances of different areas of the network and the cluster heads from the sink are significantly different. In this case, if nodes in different regions select the probability of becoming a cluster head based on their specific distance to the sink as opposed to all nodes using the same probability, further improvements to the average energy consumed by each node may be achieved. Now Eqn. (8) is a general expression for the average energy consumed by a node at a distance of d from the sink. The number of nodes in a cluster will also vary with the distance. However, if the probability of becoming a cluster head in a region with average distance of d from the sink is denoted by p, then the expected number of nodes in a cluster at this distance is again given by $N = \frac{1}{p}$. This is because for a group of adjacent sensor nodes, the probability of becoming a cluster head at the nodes do not have big difference, and we can assume them to be the same.

To obtain the optimal probability with which a sensor node at a distance of d from the sink should become a cluster head, we take the derivative of Eqn. (8) with respect to p and equate it to zero. This is given by

$$\begin{aligned} 0 = & E_{elec}k_{inter} - \frac{p_{rate}E_{amp}k_{data}L^2}{3Mp^2} + E_{amp}k_{data}d^2 \\ & + E_{amp}k_{data}d^2(1 - p_{rate})^{\frac{1}{p}} \left[\frac{\ln(1 - p_{rate})}{p} - 1 \right] (15) \end{aligned}$$

which can then be solved numerically to obtain the optimal probability for a given value of p_{rate} and the other system settings including the dimension of the sensor field L, the number of sensors M, the size of the messages k_{data} and the hardware specific constants E_{elec} and E_{amp}.

For the specific case where each sensor node always has data to report, i.e. where $p_{rate} = 1$, Eqn. (15) can be solved for a closed expression for p. This probability is given by

$$p = \sqrt{\frac{E_{amp}k_{data}L^2}{3M(E_{elec}k_{inter} + E_{amp}k_{data}d^2)}} \quad (16)$$

Note that from Eqn. (16) we can infer that sensor nodes which are closer to the sink have a higher probability of becoming cluster heads and as the distance of the nodes from the sink increases, the probability that a node becomes a cluster head decreases. As a consequence, the size of a cluster also increases as one moves away from the sink. Therefore, although nodes further away from the sink spend more energy on communicating with the sink, they are more likely to be leaf nodes, a state in which nodes consume less energy.

5 Simulation Results

In this section we use simulation results to obtain the network lifetime for various choices of the probability with which nodes become cluster heads and show that the optimal probability derived in the previous section indeed increases the network's lifetime. We also compare the lifetime of the network using our cluster head selection probability to that specified in the LEACH. Results are also presented that explore the relationship between various system parameters and the network lifetime. All simulation results reported in this section were conducted using code written in MATLAB. The simulations assume radio characteristics of $E_{elec} = 50nJ/bit$ and $E_{amp} = 100pJ/m^2$. The payload corresponding to each data reported by a sensor node is set at $k_{data} = 64Bytes$. The inter cluster communication time and the intra cluster communication time is set equal to the time to transmit $3000Bytes$. The data rate of the communication channel was 20kbps. Finally, the sink was located at position $(0, 0)$ and each round consists of one cycle.

To simulate the cluster head selection algorithm of LEACH specified in [1], cluster heads were stochastically selected in the simulations. In order to select cluster heads, at the beginning of each round each node selects a number between 0 and 1 randomly. For the nth node in the network, if this number is less than a threshold $T(n)$, the node becomes a cluster head for the current round. The threshold is set as follows

$$T(n) = \begin{cases} \frac{p}{1-p(r \bmod \frac{1}{p})} & n \in G \\ 0 & n \notin G \end{cases} \quad (17)$$

where p is the probability to become a cluster head, r is the current round, and G is the set of nodes that have not been cluster heads in the last $\frac{1}{p}$ rounds [1].

5.1 Simulation for Small Networks

We first consider the simulation results for the case of small networks. Here we consider a square sensor field with sides of length $L = 200m$, and consider three scenarios: where the number of nodes in the network M are 400, 1000 and 2000. We first consider the case where $p_{rate} = 1$. For these scenarios, the optimal probability for nodes to become a cluster head is given in Eqn. (14) and for the three network scenarios considered here, the simulation and analytic values are compared in Table (1). We note that the analytic results match quite well with the simulation results.

nodes number	$M = 400$	$M = 1000$	$M = 2000$
analysis	0.0258	0.0163	0.0115
simulation	0.0260	0.0160	0.0120

Table 1. Optimal p with different M

The simulation results in Table 1 were obtained by conducting simulations with different values of p and selecting the value of p that led to the longest network lifetime. The relationship between p and the average lifetime of sensor network as obtained through the simulations is shown in Figure 3, for the three different network settings. The lifetime of sensor nodes in Figure 3 is shown in normalized scale where the lifetimes are normalized to the longest observed lifetime during the simulations. The optimal value of p from these curves is noted in Table 1.

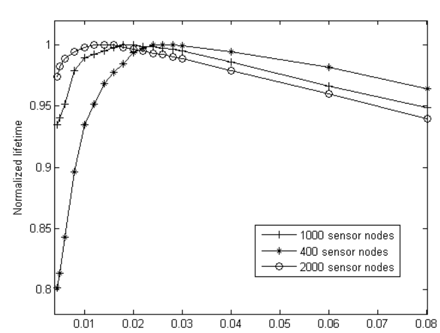

Figure 3. Network lifetime vs. p

For a given number of sensor nodes, the dimension of the sensor field may be varied to observe the impact of the area covered by the network on the optimal probability of becoming a cluster head. We now consider a network with 1000 nodes for networks covering square areas with sides of 100, 200 and 400 meters. Considering $p_{rate} = 1$, the optimal probability with which sensor nodes should become cluster heads is again given by Eqn. (14) and we compare these analytic results for this case with those obtained for simulations in Table 2. We again note the close match between the analytic and simulation results. The relation between the size of the network field and the optimal value of p as obtained through the simulations for different values of p is shown in Figure 4. From the results in Figure 4 and also from Eqns. (13) and (14) we note that as the area covered by the network becomes larger (i.e. larger L), the optimal value of p increases. This is because as the network size increases the node density decreases. Thus more sensor nodes now need to become cluster heads in order to reduce the average distance from cluster heads to the sink and reduce the power consumption.

Length of area	$L = 100$	$L = 200$	$L = 400$
analysis	0.0105	0.0163	0.0202
simulation	0.0110	0.0160	0.0200

Table 2. Optimal p for fields of different areas

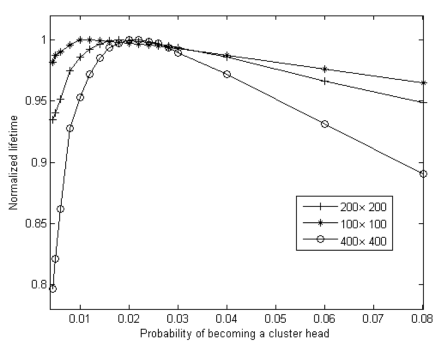

Figure 4. Network lifetime vs. p

In both Tables 1 and 2 we note that the difference between the analytic and simulation results is small. We can conclude that this approximation is valid in the case of small networks.

From Eqn. (14) we also note that if the sleep time for the leaf nodes is increased (larger T_{inter}), p will decrease.

The rate at which data is generated at the nodes is also an important factor for determining the optimal value of p. Figure 5 shows the relationship between the average energy consumption of the nodes and the probability with which nodes become cluster heads as well as the rate at which data is generated by the nodes for a network with $L = 200m$ and $M = 2000$. The optimal probability of cluster head selection increases as the rate of data arrival at each leaf node increases. The reason behind this is that since the probability that the cluster head has data to send to the sink during a cycle increases as the number of nodes in the cluster increases, reducing the cluster size (or equivalently increasing p) will reduce the energy spent by cluster heads on transmitting data to the sink.

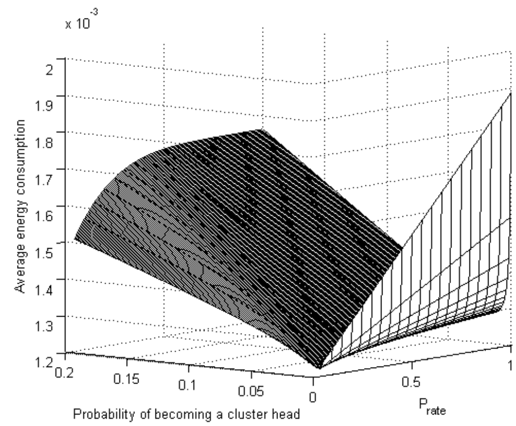

Figure 5. Average energy consumption for different cluster head probabilities and data arrival rates at leaf nodes

5.2 Simulation for Large Networks

In this section, we compare the simulation results for the normalized lifetime by small network analysis, large network analysis as well as with the basic LEACH where nodes have a 5% probability of becoming a cluster head. for the results presented in this section, the inter cluster period and the intra cluster period are set equal to the time to transmit $7000Bytes$. Also, the network considered has $L = 1000m$, and we consider three cases with 1000, 5000 and 10000 nodes. The rest of the parameters are the same as in the previous.

For comparing the results, we assume that all nodes always have data to send in each cycle($p_{rate} = 1.0$). In this case, the analysis for the small network analysis gives the optimal values of p as 0.0215, 0.0096 and 0.0068 for the 1000, 5000 and 10000 nodes cases. The normalized lifetime of the network (where the lifetime is normalized to the lifetime achieved with the optimal p values obtained through the analysis for large networks) for the the three cases is shown in Table 3. We note that in all three network settings, the lifetime of the network is the largest when the optimal p values from the large network analysis are used. We also note that the difference in the lifetimes achieved when the optimal p is calculated through the small and the large network analysis is not much for $L = 1000m$. However, the network lifetime achieved through both the small and the large network analysis is much longer than that achieved by LEACH.

In Table 4 we show the normalized lifetime for the small network analysis, the large network analysis and basic LEACH for a larger network with $L = 4000m$. For the small network analysis, the optimal values of p were 0.0223, 0.0100 and 0.0071 for the 1000, 5000 and 10000 node networks, respectively. For this network size we note that optimal p values obtained with the large network analysis achieve much longer network lifetimes as compared to LEACH. Also, the difference in the lifetime as compared to the small network analysis is much larger than in the $L = 1000m$ case.

	Small Network	Large Network	LEACH
$M = 1000$	0.990	1.000	0.891
$M = 5000$	0.996	1.000	0.872
$M = 10000$	0.993	1.000	0.813

Table 3. Comparison of lifetime $L = 1000$

	Small Network	Large Network	LEACH
$M = 1000$	0.972	1.000	0.828
$M = 5000$	0.976	1.000	0.797
$M = 10000$	0.973	1.000	0.731

Table 4. Comparison of lifetime $L = 4000$

6 Conclusion

This paper presents a framework for obtaining the optimal probability with which a node becomes a cluster head in the LEACH. We consider both small and large network scenarios and the optimal probability is obtained. Closed form solutions are also obtained for the case where all nodes have data to report in each cycle. In order to improve the scalability and energy efficiency of the MAC, this paper also presents a modified MAC protocol based on CSMA/CA with sleep and wakeup for the LEACH architecture. This new architecture is scalable, more easily implementable and reduces the hardware complexity of the sensor nodes. Simulation results are used to verify the accuracy of our calculations for obtaining the optimal probability with which nodes should become cluster heads. Our results also explore the relationship between the various system settings such as the number of sensor nodes, the dimensions of cluster field, the length of the sleep period and the rate at which data is generated by the sensor nodes on the optimal probability of cluster head selection. Our framework avoids the difficulties and overheads associated with other methods for setting the cluster head selection probability and successfully maximizes the lifetime of sensor networks.

References

[1] W. Heinzelman, A. Chandrakasan and H. Balakrishnan, "Energy-efficient communication protocols for wireless microsensor networks" *Proceedings of the Hawaii International Conference on Systems Sciences,* Jan. 2000.

[2] S. Lindsey, C. Raghavendra, K. M. Sivalingam, "Data Gathering Algorithms in Sensor Networks using Energy Metrics", *IEEE Trans. Parallel and Distributed Systems,* vol. 13, no. 9, pp. 924-935, Sept. 2002

[3] A. Manjeshwar, Q-A. Zeng, D. P. Agarwal, "An Analytical Model for Information Retrieval in Wireless Sensor Networks using Enhanced APTEEN Protocol", *IEEE Trans. Parallel and Distributed Systems,* vol. 13, no. 12, pp. 1290-1302, Dec. 2002

[4] M. J. Handy, M. Haase, D. Timmermann, "Low Energy Adaptive Clustering Hierarchy with Deterministic Cluster-Head Selection", *Proceedings of IEEE MWCN,* Stockholm, Sweden, 2002.

[5] K. Pahlavan and A. Levesque, "*Wireless Information Networks,*" New York: Wiley, 1995

[6] W. Heinzelman, "*Application-Specific Protocol Architectures for Wireless Networks*," PhD thesis, Massachusetts Inst. of Technology, June 2000.

[7] W. Ye, J. Heidemann and D. Estrin, "An Energy-efficient MAC Protocol for Wireless Sensor Networks," *Proceedings of IEEE INFOCOM,* 2001.

[8] W. Heinzelman, A. Chandrakasan and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks", *IEEE Transactions on Wireless Communication,* vol. 1, no. 4, pp. 660-670, 2002.

[9] Q. Du, M. Gunzburger, and L. Ju, "Constrained centroidal voronoi tessellations for surfaces", *SIAM Journal on Scientific Computing 24,* pp. 1488-1506, 2003.