

Applying an SOM Neural Network to Increase the Lifetime of Battery-Operated Wireless Sensor Networks

Mario Cordina and Carl James Debono
*University of Malta
Malta*

1. Introduction

Wireless sensor networks have garnered significant attention in recent years. According to (The Mobile Internet, 2004), more than half a billion nodes will be shipped for wireless sensor applications in 2010, for an end user market worth at least \$7 billion.

Wireless sensor networks are one of the first real-world examples of pervasive computing, the notion that small, smart, computing and cheap sensing devices will eventually permeate the environment (Bulusu & Jha, 2005). The combination of distributed sensing, low power processors and wireless communication enables such technology to be used in a wide array of applications such as habitat monitoring and environment monitoring, military solutions, such as battlefield surveillance, and commercial applications, such as monitoring material fatigue and managing inventory.

A wireless sensor network consists of hundreds or thousands of low-power, low-cost multi-functioning sensor nodes operating in an unattended environment with a limited supply of energy. The latter is one of the main constraints of each sensor node together with the limited processing power. These limitations, coupled with the deployment of a large number of sensor nodes, pose a number of challenges to the design and management of these networks, requiring energy-awareness at all layers of the networking protocol stack.

The issues related to the physical and link layers are generally common for all sensor applications and therefore research in these areas focused on system-level energy awareness such as dynamic voltage scaling (Heinzelman et al, 2000a), radio communication hardware (Min et al, 2000), low duty-cycle issues (Woo & Culler, 2001), system partitioning (Ye et al, 2002) and energy-aware MAC protocols (Shih et al, 2001). At the network layer, energy efficient route setup protocols are necessary to reliably relay data from the sensor nodes to the sink whilst maximising the lifetime of the network. This chapter will focus on such a solution based on sensor node clustering, whereby the topology is decided through an SOM neural network.

2. Related Work

Since wireless sensor networks often consist of a large number of sensor nodes which have to be networked together, conventional techniques such as the direct transmission protocol have to be avoided because the energy loss incurred can be quite large, depending on the location of the sensor node relative to the base station. Furthermore, using a conventional multi-hop routing protocol such as Minimum Transmission Energy (MTE) (Ettus, 1998; Shepard, 1996) will also result in an equally undesirable effect. This is due to the fact that in MTE, the intermediate nodes are chosen in such a way that the sum of squared distances (and thus the total transmit energy, assuming d^2 power loss) is minimised. Hence for the configuration shown in Figure 1, node A would transmit to node C through node B if and only if:

$$E_{\text{transmit}}(d = d_{AB}) + E_{\text{transmit}}(d = d_{BC}) < E_{\text{transmit}}(d = d_{AC}) \quad (1)$$

or assuming a $\frac{1}{d^2}$ attenuation model,

$$d_{AB}^2 + d_{BC}^2 < d_{AC}^2 \quad (2)$$

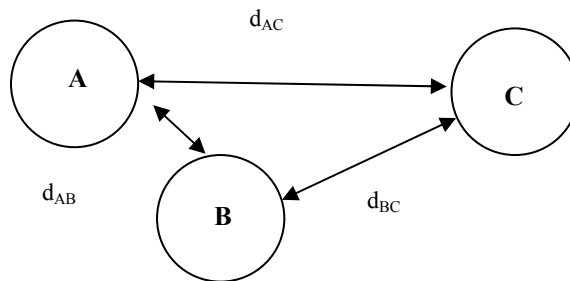


Fig. 1. Minimum Transmission Energy

Thus in MTE, the nodes closest to the base station will rapidly drain their energy resources since they are involved in routing of a large number of messages (on behalf of other nodes) to the base station.

In order to minimise these problems, various routing protocols have been proposed for wireless sensor networks. An in-depth survey of such protocols is presented in (Akkaya & Younis, 2005). Here, Akkaya and Younis classify routing protocols into 3 main categories namely: Location-based protocols, Data-centric protocols and Cluster-based protocols.

Location-based protocols

These protocols utilise the position information to relay data to the desired regions rather than the whole network. An example of location-based protocols is Geographic Adaptive Fidelity (GAF) (Xu et al, 2001).

Data centric protocols

These protocols depend on labelling of the desired data thereby eliminating many redundant transmissions. Two well-known network layer protocols based on data centric

routing are Sensor Protocols for Information via Negotiation (SPIN) (Heinzelman et al, 1999) and directed diffusion (Intanagonwiwat et al, 2000). Although the data centric routing approach provides a reliable and robust solution to wireless sensor networks, there are still some shortcomings associated with it since both SPIN and directed diffusion suffer from the amount of overhead energy spent in activities such as advertising, requesting and gradient setup. Furthermore, the excessive time spent in such activities might not suit applications that require the sensor nodes to respond quickly to an emergency situation.

Cluster-based protocols

In cluster-based routing protocols, nodes are grouped into clusters and each cluster head node collects, processes and forwards the data from all the sensor nodes within its cluster to the base station. The application of clustering-based protocols reduces energy dissipation by selecting optimal cluster heads, thereby reducing transmission distance, and by reducing the amount of information that needs to be transmitted through data aggregation. Research in recent years focused on developing algorithms that select optimal cluster heads in order to reduce the energy dissipation thereby increasing the lifetime of the system.

Cluster-based routing algorithms are generally divided into two categories; centralised and distributed algorithms. One of the most popular cluster-based distributed algorithms is Low-Energy Adaptive Clustering Hierarchy (LEACH) (Heinzelman et al, 2000b). The operation of LEACH is organized in rounds, where each round is composed of a cluster setup phase and a data transmission phase. During the cluster setup phase, nodes organize themselves into clusters with one node serving as the cluster head in each cluster. The selection of a cluster head is decided locally within each node, making LEACH a completely distributed algorithm requiring no global knowledge of the network. To become a cluster head, each node n selects a random number between 0 and 1 and if the number is less than the threshold $T(n)$ the node is elected as a cluster head for the current round. The threshold $T(n)$ is given by:

$$T(n) = \begin{cases} \frac{P}{1 - P \left(r \bmod \frac{1}{P} \right)} & \text{if } n \in G \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

where, P is the cluster head probability, r is the number of the current round and G is the set of nodes that have not been cluster-heads in the last $\frac{1}{P}$.

During the transmission phase, the self elected cluster heads collect data from nodes within their respective clusters and apply data fusion before forwarding the fused data directly to the base station. At the end of a given round, a new set of nodes become cluster heads for the subsequent round. This cluster head rotation mechanism coupled with data fusion in the cluster head results in LEACH achieving a factor of 7 reduction in energy dissipation compared to direct communication and a factor of 4 - 8 reduction in energy when compared to the minimum transmission energy routing protocol (Akkaya & Younis 2005).

Although LEACH is normally considered as a benchmark protocol, the algorithm used in the selection of cluster heads has a number of disadvantages:

1. Since each node uses probability techniques to decide whether or not to become a cluster head, there might be cases where two cluster heads are in close proximity of each other increasing the overall energy dissipation in the network.
2. The number of cluster head nodes generated is not fixed and so in some rounds it may be more or less than the optimal value.
3. The nodes selected can be located in areas which are not densely populated resulting in further energy expenditure to transmit data to that cluster head.
4. Each node has to calculate the threshold and generate the random number in each round, consuming CPU cycles.
5. LEACH stochastic cluster head selection is prone to lead to energy imbalance in the network. This results in nodes dying at an early stage reducing the network lifetime significantly.

In order to tackle some of the disadvantages found in LEACH, the authors of (Heinzelman et al, 2002) proposed a centralized version called LEACH-C. Unlike LEACH, where nodes self-configure into clusters, LEACH-C uses the base station for cluster formation. During the cluster setup phase, each node sends information about its current location, using a GPS receiver, and energy level to the base station. The energy level is used by the base station to ensure that only the nodes with relatively high energy are participating in the cluster head selection. Using these high energy nodes, the base station uses the simulated annealing algorithm (Maruta & Ishibuchi, 1994) to find k optimal cluster heads in order to minimize the amount of energy dissipation required by the other nodes to transmit their data to the cluster head.

Although the data transmission phase of LEACH-C is identical to that of LEACH, results indicate a 40% increase in data per unit energy over LEACH. The authors of (Heinzelman et al, 2002) cite two key reasons for this improvement:

1. The base station utilises its global knowledge of the network to produce better clusters which require less energy for data transmission.
2. The number of cluster heads in each round of LEACH-C equals a predetermined optimal value, whereas for LEACH the number of cluster heads varies from round to round due to the lack of global coordination among nodes.

Nevertheless, this improvement comes at the expense of having all the sensors equipped with GPS receivers which apart from increasing the cost of each sensor, introduces further energy dissipation.

3. Algorithm Design

3.1 Requirements and Assumptions

In order to find a good compromise between the centralised and distributed approaches, we will now present an SOM neural network-based clustering algorithm. In this solution, the sensor nodes collect network topological information, through localised interactions, which is then transferred to the base station through a low energy cost network initialisation phase. As opposed to other centralised algorithms such as LEACH-C, the initialisation phase allows the base station to have updated network topology awareness without the use of GPS-assisted sensors and without high energy cost procedures.

The design of the algorithm needs to achieve the following targets:

- **Ease of deployment:** Sensor networks may contain hundreds of nodes and may need to be deployed in remote and dangerous environments. Thus nodes must be small, cheap and able to self-configure with limited global control to setup or maintain the network.
- **System lifetime:** Sensor nodes must have low energy consumption in order to allow the network to operate for as long a period as possible. Moreover, the energy balancing mechanisms help in extending the system’s lifetime.
- **Latency:** Data from sensor networks is typically time-sensitive, so it is important to receive the data in a timely manner. Long delays due to processing or communication may be unacceptable.

In order to reach these goals, the algorithm mechanisms described below and illustrated in Figure 2 can be applied.

System Lifetime: To improve system lifetime, mechanisms need to be developed to minimise energy dissipation and improve energy balancing between the nodes. Such mechanisms include: Data reduction algorithms based on classic adaptive filters, cluster head separation, neural assisted cluster head election, cluster head rotation and load balancing cost functions.

Latency: Network latency increases as number of nodes between source and destination increases. To reduce latency, the single hop mode of operation can be used for intra-cluster communication, thereby allowing the base station to receive data in a timely manner. Furthermore, the use of data reduction algorithms discussed later on, allow the base station to predict sensor measurements itself thus eliminating any latency issues.

Ease of deployment: The use of a homogenous network structure whereby all the sensor nodes are identical in terms of battery energy and hardware complexity allows for a fast and easy deployment. Moreover, the algorithm can be designed in such a way that nodes self-configure with limited global control to setup or maintain the network.

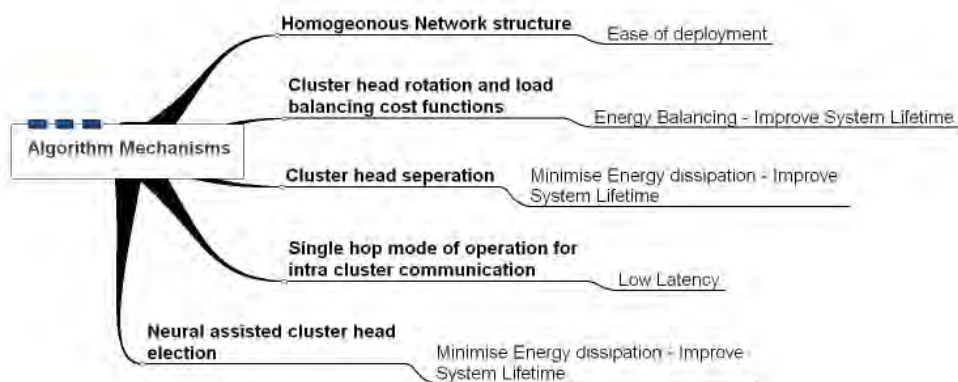


Fig. 2. Algorithm Mechanisms

Prior to a detailed description of the algorithm, it is worth noting the following general assumptions:

- N sensors are uniformly dispersed within a square field of size M x M.
- All sensors and base station are stationary after deployment.

- The intra-cluster communication is based on the single-hop mode of operation.
- Communication is symmetric and a sensor can compute the approximate distance based on the received signal strength if the transmission power is given. If the transmission power is not known, the received signal strength is used as a distance metric.
- All sensors are of equal significance having equal energy and hardware complexity – homogenous network.
- All sensors are location-unaware.

3.2 Radio Propagation Model

The transmission of data from the sensors uses air as a medium and therefore we need a model to estimate the signal levels reaching the receivers. The standard radio model used in wireless sensor networks is based on the fact that the propagation of electromagnetic waves can be modelled as a decaying power law function of the distance between the transmitter and the receiver. Furthermore, if there is no direct line-of-sight path between the transmitter and the receiver, the electromagnetic wave will undergo reflection, diffraction and scattering off objects in the environment. This will result in electromagnetic waves arriving at the receiver from different paths and at different times, leading to multi-path fading, which again can be roughly modelled as a power law function of the distance between the transmitter and the receiver ([Heinzelman et al, 2000b](#)).

The radio model commonly adopted to model propagation in wireless sensor networks uses both the free space model and the multi-path fading model, depending on the distance between the transmitter and the receiver. If the distance between the transmitter and a receiver is less than a certain crossover distance $d_{crossover}$, the Friss free space model is used whereas if the distance is greater than $d_{crossover}$, the two-ray ground propagation model is used ([Heinzelman et al, 2000b](#)). The crossover distance is defined as follows ([Heinzelman et al, 2000b](#)):

$$d_{crossover} = \frac{4\pi\sqrt{L}h_T h_R}{\lambda} \quad (4)$$

where,

$L \geq 1$ is the system loss factor and is not related to propagation,

h_T is the height of the transmitting antenna above ground level in metres,

h_R is the height of the receiving antenna above ground level in metres,

λ is the wavelength of the carrier signal in metres.

When the transmitter and receiver have direct line-of-sight communication, which will only occur if the transmitter and receiver are close to each other (i.e. $d \leq d_{crossover}$), the transmit power is attenuated according to the Friss free space equation as follows ([Rappaport, 1996](#)):

$$P_r(d) = \frac{P_t G_t G_r \lambda^2}{(4\pi d)^2 L} \quad (5)$$

where,

$P_r(d)$ is the received power given a transmitter-receiver separation of d in metres,

P_t is the transmit power,

G_t is the gain of the transmitting antenna,

G_r is the gain of the receiving antenna,

d is the distance between the transmitter and the receiver in metres.

If the distance is greater than $d_{crossover}$, the transmit power is attenuated according to the two-ray ground propagation equation as follows ([Heinzelman et al, 2000b](#)):

$$P_r(d) = \frac{P_t G_t G_r h_t^2 h_r^2}{d^4} \quad (6)$$

In this case, the received signal comes from both the direct path and a ground-reflection path leading to a higher transmit power attenuation rate proportional to d^4 .

Although the standard radio model is widely used in modelling and simulation of wireless sensor networks, it is rather simplistic. As the credibility of high level protocol simulation results depends on the accuracy of the physical layer model ([Kotz et al, 2004](#)), more accurate radio models are required. A review of studies carried out in this area reveals a general lack of near-ground channel measurements, as the vast majority of the studies place antennas at heights which are greater than one metre. However, recently studies were carried out in this area and radio models based on field measurements were specifically developed for wireless sensor network. In ([Fanimokun & Frolik, 2003](#)), Fanimokun and Frolik present specific models for near ground wireless sensor networks operating in the 915 MHz ISM band based on the single slope log normal shadowing model. Similarly in ([Molina-Garcia-Pardo et al, 2005](#)), the authors present models for an operational frequency of 868 MHz based on the two slope log normal models. Although these models are more accurate than the standard radio model discussed earlier, they are specific to the environment in which the measurements were taken and therefore they cannot be reliably applied to different scenarios.

3.3 Energy Model

The standard energy consumption model used in wireless sensor networks is that developed by ([Heinzelman et al, 2000b](#)). Heinzelmann adopts a simple energy model where the transmitter dissipates energy to run the radio electronics and the power amplifier, whilst the receiver dissipates energy to run the electronics as shown in Figure 3.

Furthermore, this energy model makes use of the standard propagation model whereby the power attenuation depends on the distance between the transmitter and receiver. Power control is used to invert this loss by adjusting the power amplifier to ensure that a certain signal power level reaches the receiver.

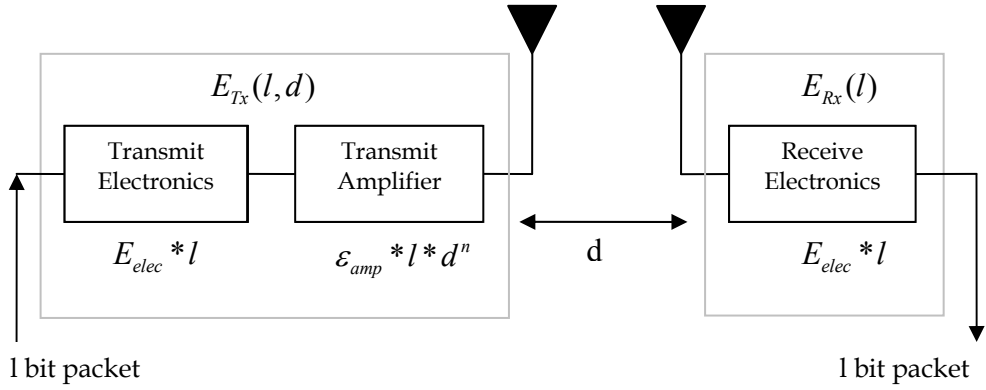


Fig. 3. Radio energy dissipation model (Heinzelman et al, 2000b)

Thus, to transmit an l -bit message a distance d , the radio expends:

$$E_{TX}(l, d) = E_{TX-Elec}(l) + E_{TX-amp}(l, d) \quad (7)$$

This is equivalent to:

$$E_{TX}(l, d) = \begin{cases} lE_{Elec}(l) + l\epsilon_{fs}d^2 & ; d < d_{crossover} \\ lE_{Elec}(l) + l\epsilon_{mp}d^4 & ; d \geq d_{crossover} \end{cases} \quad (8)$$

And to receive this message, the radio expends:

$$E_{Rx}(l) = E_{Rx-elec}(l) \quad (9)$$

$$E_{Rx}(l) = lE_{Elec}(l) \quad (10)$$

The electronics energy E_{Elec} depends on factors such as digital coding, modulation and filtering of the signal before it is sent to the transmit amplifier. The typical value used for E_{Elec} is 50 nJ/bit (Heinzelman et al, 2000b). The parameters ϵ_{fs} and ϵ_{mp} depend on the receiver sensitivity and noise figure, as the transmit power needs to be adjusted so that the power at the receiver is above a certain threshold $P_{Threshold}$ (Heinzelman et al, 2000b).

3.4 Algorithm Details

The general operation of the clustering algorithm is divided into three phases as shown in Figure 4.

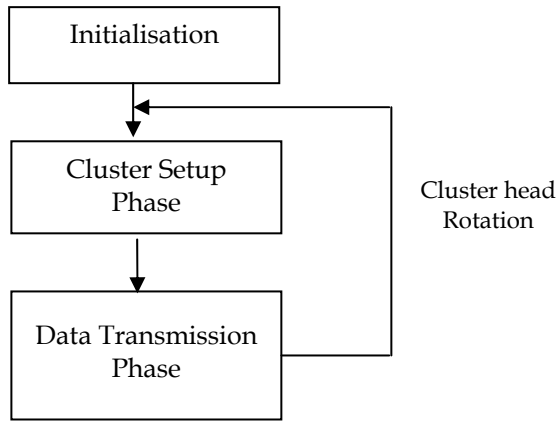


Fig. 4. General operation of the Algorithm

Similar to LEACH, the operation of the algorithm is divided into rounds. Each round begins with a cluster setup phase where the clusters are organised, followed by a data transmission phase during which data is transferred from the nodes to the cluster head which aggregates all the data received together with its own data and forwards it to the base station.

Initialisation Phase

During the initialisation phase, the base station calculates important network parameters and collects topological information from the nodes through localised interaction. This phase needs to consume very little energy and therefore it must be designed to use small control packets and minimise the number of transmissions. The operation of the initialisation phase is as follows:

1. The base station calculates the optimal number of cluster heads K_{opt} and the minimum cluster head separation distance R_{sep} .
2. The base station broadcasts a "Network Initialisation" message to all nodes in the network at a fixed power $P_{Broadcast}$ which includes the minimum separation distance R_{sep} .
3. All the nodes in the network receive this broadcast and use it as a 'beacon' signal to calculate their distance from the base station d_{TOBS} by considering the received signal strength as a distance metric.
4. Each node then broadcasts a "Neighbourhood Search" message including their node ID within a range R_{sep} by setting the transmit power to P_{Search} .
5. All the nodes in the network listen for the "Neighbourhood Search" messages and record all the received messages, storing the Node ID and received signal strength.
6. Following the "Neighbourhood Search" procedure, each node uses the recorded "Neighbourhood Search" messages to calculate its node centrality and concentration as in (Gupta et al, 2005). These parameters together with the list of neighbouring nodes and the distance metric to each node are then transmitted to the base station.

Optimum number of cluster heads

The optimal number of cluster heads, K_{opt} , can be analytically determined using the computation and communication energy models as in LEACH (Heinzelman et al, 2000b). The authors of (Heinzelman et al, 2000b), considered the case where the cluster head aggregates all the received data into a single packet by performing data aggregation. However, this is not realistic and the case where the cluster head compresses all the received data into a packet of size ρl is considered.

Assume that there are N nodes distributed uniformly in an $M \times M$ region. If there are k clusters, each cluster has an average of N/k nodes (one cluster head and $(N/k) - 1$ non-cluster head nodes). Now, let $\rho \in [0,1]$ be the compression factor, l be the length of data packets transmitted from the node to the cluster head, P be the number of non-cluster head nodes in a cluster and L_M be the length of the compressed data packet transmitted from the cluster head to the base station.

Then,

$$L_M = (P - (P - 1)\rho)l \quad (11)$$

Since the base station is typically far from the nodes, it is assumed that the energy dissipation of the cluster head follows the multi-path model (d^4 power loss). Thus, the energy dissipated in the cluster head during a single frame is given by:

$$E_{CH} = \left(\frac{N}{k} - 1\right)lE_{elec} + L_M \left(E_{elec} + \varepsilon_{mp} d_{TOBS}^4\right) \quad (12)$$

where E_{elec} is the electronics energy taken to be 50nJ/bit (Heinzelman et al, 2000b), ε_{mp} depends on the receiver sensitivity and noise figure, and d_{TOBS} is the distance from the cluster head node to the base station.

Each non-cluster head node transmits its data to the cluster head once during a frame. As the distance to the cluster head is small, it is assumed that the energy dissipation of the non-cluster head node follows the Friss free space model (d^2 power loss). Thus the energy dissipated in each non-cluster head node is:

$$E_{non-CH} = lE_{elec} + l\varepsilon_{fs}d_{TOCH}^2 \quad (13)$$

where ε_{fs} depends on the receiver sensitivity and noise figure and d_{TOCH} is the distance from the non-cluster head node to the cluster head. As shown in (Heinzelman et al, 2000b),

$$E[d_{TOCH}^2] = \frac{M^2}{2\pi k} \quad (14)$$

Therefore, the energy dissipation of the non-cluster heads is given by:

$$E_{non-CH} = lE_{elec} + l\varepsilon_{fs} \frac{M^2}{2\pi k} \quad (15)$$

Furthermore, the total energy dissipated in a cluster during a frame is:

$$E_{cluster} = E_{CH} + \left(\frac{N}{k} - 1\right)E_{non-CH} \approx E_{CH} + \frac{N}{k}E_{non-CH} \quad (16)$$

Hence the total energy dissipated in a frame is given by:

$$E_{total} = kE_{cluster} = k \left[E_{CH} + \frac{N}{k}E_{non-CH} \right] \quad (17)$$

Substituting (12) and (13) in (17) yields:

$$E_{total} = nl \left[2E_{elec} + \varepsilon_{fs} \frac{M^2}{2\pi k} \right] - klE_{elec} + kL_M \left[E_{elec} + \varepsilon_{mp}d_{TOBS}^4 \right] \quad (18)$$

The optimal number of cluster heads is found by setting the derivative of E_{total} with respect to k to zero, yielding:

$$k_{opt} = \sqrt{\frac{nl\varepsilon_{fs}M^2}{2\pi \left[\rho l E_{elec} - 2lE_{elec} - l\varepsilon_{mp}d_{TOBS}^4 + \rho l\varepsilon_{mp}d_{TOBS}^4 + \left[E_{elec} + \varepsilon_{mp}d_{TOBS}^4 \right] \rho l \right]}} \quad (19)$$

Minimum cluster head separation distance

In order to reduce the energy dissipation of the sensor nodes, cluster heads are spread across the whole field by considering a minimum cluster head separation distance, R_{sep} . This ensures that no two cluster heads are in close proximity of each other, thereby increasing network lifetime.

Consider K_{opt} cluster heads distributed uniformly in an $M \times M$ region. The minimum cluster head separation distance R_{sep} is given by:

$$R_{sep} = \sqrt{\frac{M^2}{\pi k_{opt}}} \quad (20)$$

Determining $P_{Broadcast}$ and P_{Search}

With reference to equations (6), (7), (8) and assuming the following radio parameters: $f = 914\text{MHz}$, $\lambda = 0.328\text{m}$, $G_T = G_R = 1$, $L = 1$, $h_T = h_R = 1.5\text{m}$ (Heinzelman et al, 2000b), the crossover distance is given by:

$$d_{\text{crossover}} = \frac{4\pi\sqrt{L}h_T h_R}{\lambda} = 86.2m \quad (21)$$

Thus,

$$P_R = \begin{cases} \left[\frac{G_T G_R \lambda^2}{(4\pi)^2} \right] \frac{P_T}{d^2} = 6.82 * 10^{-4} \frac{P_T}{d^2} & \text{if } d < d_{\text{crossover}} \\ G_T G_R h_T h_R \frac{P_T}{d^4} = 5.1 * \frac{P_T}{d^4} & \text{otherwise} \end{cases} \quad (22)$$

Assuming that the receiver's thermal noise floor is 99dBm, the receiver noise figure is 17dB and assuming that an SNR of at least 30dB is required to receive the signal with no errors, then the minimum received signal level for successful reception is given by (Heinzelman et al, 2000b): $P_R \geq -99 + 17 + 30 = -52\text{dBm} \rightarrow 6.3\mu\text{W}$.

Substituting P_R in (22) yields,

$$P_T = \begin{cases} 9.24 * 10^{-3} d^2 & \text{if } d < d_{\text{crossover}} \\ 1.24 * 10^{-6} d^4 & \text{otherwise} \end{cases} \quad (23)$$

In general, to determine P_{Search} we consider $d < d_{\text{crossover}}$ whereas to determine $P_{\text{Broadcast}}$ we consider $d \geq d_{\text{crossover}}$.

Cluster Setup Phase

During the cluster setup phase, the base station has to elect K_{opt} cluster heads. This decision is based on four parameters, namely:

- *Node Residual Energy*: The elected node must have a high residual energy as the cluster heads are involved in high energy consumption operations. These nodes have to aggregate the data and transmit larger packets to the base station.
- *Centrality*: The node centrality defines a value which classifies the node importance based on how central the node is to the cluster. The elected node needs to have a low centrality value indicating that the node is central in relation to the surrounding nodes thereby reducing the energy dissipation between the nodes and the cluster head.
- *Concentration*: This defines the number of nodes present in the vicinity. The elected node must have a high node concentration thereby ensuring that cluster heads are elected where they are mostly required.
- *Cluster head frequency*: In order to avoid electing the same node repeatedly, the base station has to keep a record of the number of times each node was elected as cluster head.

The election of cluster heads is carried out using an SOM neural network (Kohonen, 1995). The neural network uses the above mentioned parameters to partition the nodes according to their cluster head quality. From this we define high quality cluster heads as those sensor nodes that present high residual energy, low centrality, high concentration and low cluster head frequency.

A 4 input - 16 output neuron SOM topology is adequate to solve this cluster head selection problem (Cordina & Debono, 2008). The network is used to partition data into 16 categories

depending on the cluster head quality as shown in Figure 5. The weights $W_{i,j}$ are initially set such that neuron Y_1 corresponds to high quality nodes whereas neuron Y_{16} corresponds to low quality nodes. These weights are then fine tuned using the standard SOM training methodology based on 10,000 normalised test vectors.

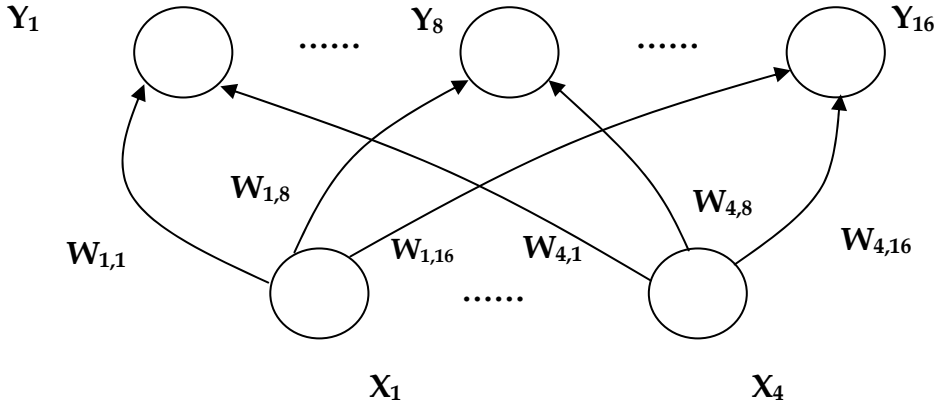


Fig. 5. Self Organising Map structure

The general operation of the cluster head election mechanism is shown in Figure 6.

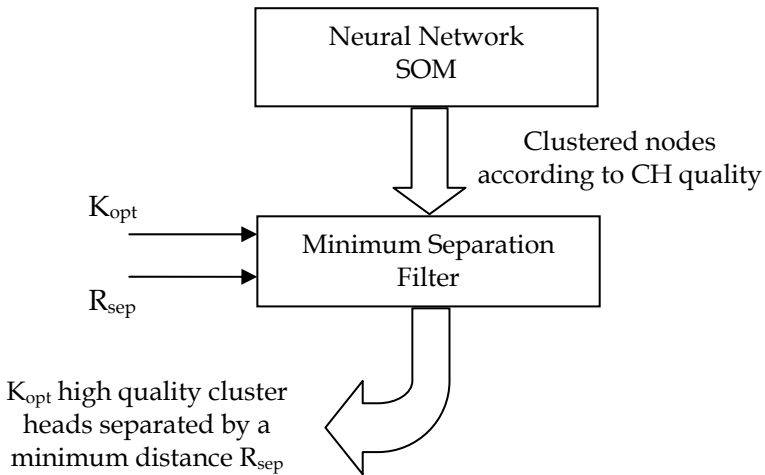


Fig. 6. Cluster head election mechanism

The output from the neural network is a list of clustered nodes sorted according to their cluster head quality. This list is then applied to a minimum separation filter which ensures that a minimum separation distance is guaranteed between the elected cluster heads. This

yields K_{opt} high quality cluster heads which are separated by the defined minimum distance R_{sep} .

The operation of the cluster setup phase is as follows:

1. During the initialisation phase, the base station collects all the information sent from the nodes and compiles a table as shown in Table 1. Note that the initial energy of each node is predefined and the cluster head frequency, which defines the number of times the node is elected as cluster head, is initially set to zero.

Node ID	Centrality	Concentration	Energy	CH Frequency	Distance to BS	Neighbours
1	0.34	1	0.5	0	80.5	54
2	0.532	3	0.5	0	75.2	5, 10, 14
3	0.784	2	0.5	0	97.6	65, 88
...
...

Table 1. Information collected by the base station

2. The Centrality, Concentration, Energy and CH frequency for all the nodes are extracted from Table 1, normalised (Doherty et al, 2007) and applied to the neural network to cluster the nodes according to their cluster head quality. Following the clustering process, the filtering algorithm is applied to the output of the neural network.
3. Using Table 1, the base station computes the cost function parameters d_{g_min} and d_{g_max} on the chosen CHs as follows

$$d_{g_min} = \min(d_{TOBS})$$

$$d_{g_max} = \max(d_{TOBS})$$

4. The base station informs the selected cluster head nodes for the current round and also includes the cost function parameters.
5. The selected cluster heads broadcast an "Advert CH" message to all the nodes in the network at a fixed power $P_{Broadcast}$ including their Node ID and cost function parameters.
6. Nodes in the network receive the "Advert CH" messages and compute the cost for each cluster head. Various cost functions can be used, two of which are shown here:

Standard cost function based on received signal strength – In most of the algorithms, nodes choose the cluster heads based on the received signal strength in order to lower their energy consumption. However as discussed in (Ye et al, 2005), this approach may lead cluster head nodes to exhaust their energy rapidly during the data transmission phase as cluster head nodes which are further away from the base station have a higher energy consumption compared to those being in the vicinity of the base station.

Advanced cost function based on a weighted distance-energy metric – This cost function presented in (Cordina & Debono, 2009) considers both the distance metric and the energy of the cluster head node relative to that of the node. Thus, the cost of a cluster head node which is far

away from the base station is offset according to the energy of the cluster head node relative to that of the node. The cost function is given by:

$$Cost(j, i) = w_1 * f(d(P_j, CH_i)) + (1 - w_1) * g(d(CH_i, BS)) - (1 - w_2) * \frac{E_{CH}}{E_{Node}} \quad (24)$$

$$f = \frac{d(P_j, CH_i)}{d_{f_max}}; g = \frac{d(CH_i, BS) - d_{g_min}}{d_{g_max} - d_{g_min}}$$

where, E_{CH} and E_{Node} are the cluster head and node residual energy respectively, w_1 and w_2 are the distance and energy cost function weights respectively, d_{g_min} and d_{g_max} are cost function distance parameters described earlier and $d_{f_max} = Ex[\max\{d(P_j, CH_i)\}]$ is a normalising factor for the distance between the node P_j and the cluster head CH_i .

7. Each node selects the cluster head which has the minimum cost and sends a "Join Request" message to the selected cluster head node including Node ID, cost and status of the exclusive flag. The latter is used by the node to inform the cluster head that it has received only one "Advert CH" message, this indicates that the node is on the fringe of the network. The exclusive flag is then used to prioritise between the nodes.
8. Cluster heads receive the "Join Request" messages and notify the nodes of their acceptance.

The nodes whose energy level falls below the Critical Energy Level whilst in the cluster setup phase send a "Node dead" message directly to the base station. Similarly, elected cluster head nodes whose energy level falls below the Critical Energy Level during the cluster setup phase, send a "Cluster Head dead" message directly to the base station. This allows the base station to update the status of the nodes in the network during the cluster setup phase. If the base station detects that a cluster head is 'dead', it marks its status 'dead' and broadcasts a re-cluster message to the whole network indicating that a re-cluster is necessary. On receiving a 'Re-clustering' message, all the nodes disassociate themselves from their cluster head and the cluster setup phase is re-started.

Data Transmission Phase and cluster head rotation

During the data transmission phase, the nodes perform measurements and append status and residual energy information to each packet before transmitting it to their cluster head. The cluster heads receive this data, append their status and residual energy information, compress it and forward the resulting packet to the base station. This allows the base station to update the energy level and status of all the nodes in the network. Whenever the base station detects that a cluster head is 'dead', it broadcasts the re-cluster message.

As there is a cost in terms of time and energy associated with the cluster setup phase, the data transmission phase should be long when compared to this phase to reduce the effect of the overhead incurred during cluster formation on the overall performance. On the other hand, as the energy at each node is limited, running the data transmission phase for too long drains the energy of the cluster head node and curtails communication between the non-cluster head nodes that still have energy and the base station. For this reason, a cluster head rotation mechanism is employed to balance energy between the nodes. In this algorithm, a cluster head rotation is triggered when the cluster head residual energy in the current round falls below the residual energy in the previous round by some percentage hysteresis. On a

cluster setup phase re-start, the base station has an updated status of the network and recalculates the optimal number of cluster heads K_{opt} to ensure optimality throughout the lifetime of the network. All these energy overheads related to the initialisation phase, cluster setup phase, associated re-clustering procedures and the appending of status/residual energy to the transmitted packets must be considered in the simulations.

3.5 Results

Prior to fully testing the algorithm, the two SOM neural network parameters, namely the learning rate and sigma, need to be optimised. Simulations are initially carried out on a fixed network in order to determine these optimal parameters. The clustering performance of the neural network is assessed by computing the overall cluster quality index as defined in (He et al, 2003). With reference to Figure 7, the optimal SOM neural network parameters are: learning rate=0.1 and sigma=0.2.

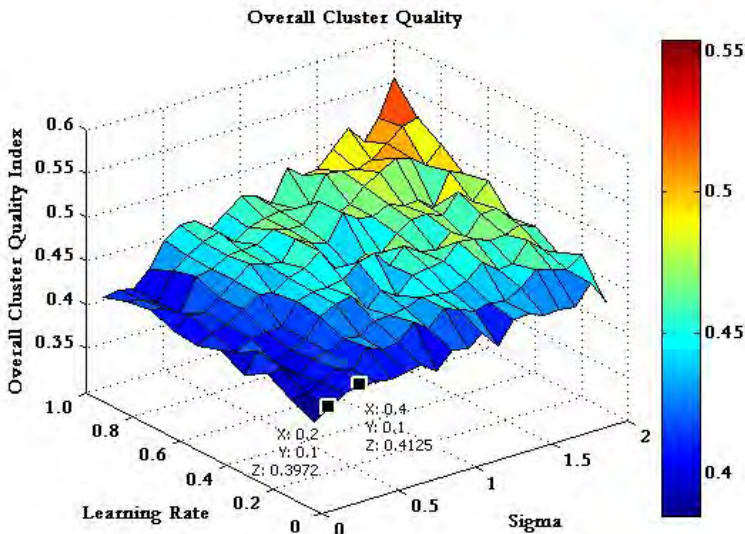


Fig. 7. Overall cluster quality index

Using these optimal SOM neural network parameters, the full algorithm can be tested in a wireless sensor network simulator, in this example the simulator is implemented in MATLAB®. Simulations were performed on 100 nodes uniformly dispersed in a 100m by 100m field with the base station located at coordinates (150m, 50m). The energy parameters used in this simulation are tabulated in Table 2.

Initial Battery energy	0.5 Joules
Energy model parameter, ϵ_{fs}	1×10^{-11}
Energy model parameter, ϵ_{mp}	1.3×10^{-15}
Electronics Energy, E_{elec}	50nJ/bit
Data packet length	4000bits
Control packet length	200 bits

Table 2. Assumed energy parameters

The performance of the algorithm can be assessed along three metrics. These are (1) *First Node Dies (FND)* which defines the time taken for the first node to die, (2) *Half Nodes Alive (HNA)* which defines when half the network has died, and (3) *Transmitted Packets (TP)* which gives the total number of transmitted packets by the sensor nodes at FND. To evaluate the validity and performance of the algorithm, simulations are carried out over 50 randomly generated 100 node networks. The improvement over LEACH in terms of mean and standard deviation values of the performance metrics is shown in Table 3.

Metric	Standard Cost Function		Advanced Cost Function	
	Mean	Std. Deviation	Mean	Std. Deviation
FND	51.1%	4.8%	65.2%	4.6%
HNA	34.%	0.9%	32.8%	0.8%
TP	27.6%	4%	36.1%	3.3%

Table 3. Performance Metrics

Algorithm	Improvement over LEACH
EEPSC (Zahmati et al, 2007)	45%
LEACH-C (Heinzelman, 2002)	62.5%
LEACH-Deterministic (Handy et al, 2003)	30%
This SOM-based Algorithm	65.2%

Table 4. Performance Improvement of various algorithms

These results show that the algorithm is capable of creating better clusters, increasing the energy balancing between the nodes whilst lowering the rate of energy dissipation. This in turn results in a significant improvement in the FND and HNA, leading to an increase in the number of packets received by the base station. Furthermore, the use of a cost function based on a weighted distance-energy metric (with w_1 and w_2 set to 0.9) further optimises the energy dissipation and energy distribution between the nodes leading to an additional improvement in the network lifetime, with an increase of up to 65.2% in FND over LEACH

as shown in Table 3, and a reduction in the FND to HNA transition time. A graphical representation comparing LEACH with the proposed algorithm is illustrated in Figure 8 while the superiority of this algorithm, based on an SOM neural network, with respect to other solutions is shown in Table 4.

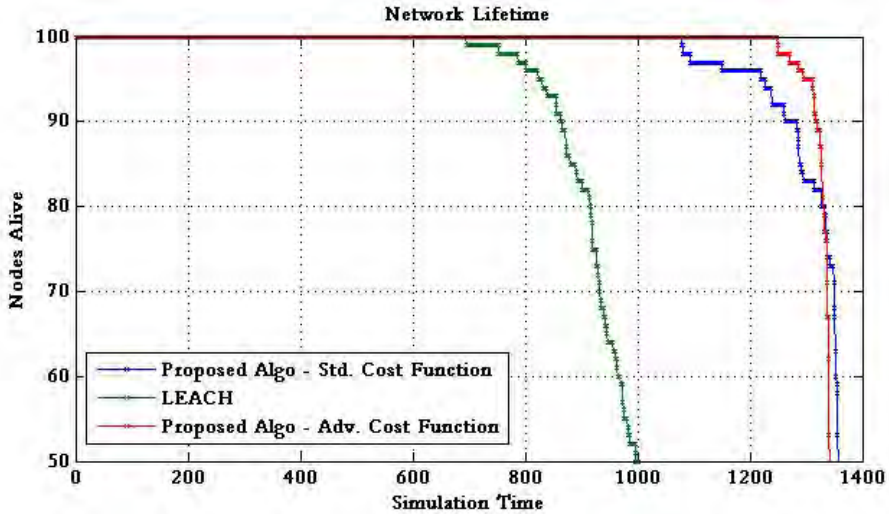


Fig. 8. Network Lifetime

4. Conclusion

We have presented a cluster-based routing algorithm developed using an SOM neural network architecture together with a cost function based on a weighted distance-energy metric. Simulation results show the efficacy of the algorithm as it manages to improve the system lifetime by up to 65.2% when compared to LEACH. Furthermore, by optimising the distribution of the energy between the nodes, a reduced FND to HNA transition time is achieved thereby improving the network quality. This solution makes wireless sensor networks more attractive as they can remain operational for a longer period of time.

5. References

- The Mobile Internet (2004), *Wireless Sensor networking: \$7 billion market by 2010*. [Online]. Available: http://findarticles.com/p/articles/mi_m0NZB/is_4_6/ai_n6054921
- Akkaya K.; Younis M. (2005). A survey of Routing Protocols for Wireless Sensor Networks *Elsevier Ad Hoc Network Journal*, Vol. 3, pp. 325-349, 2005
- Bulusu, N.; Jha, S. (2005). *Wireless sensor networks - A System Perspective*, Artech House, Boston, 2005
- Cordina, M.; Debono, C.J. (2008). Increasing Wireless Sensor Network Lifetime through the Application of SOM Neural Networks, *Proceedings of the 2008 3rd International Symposium on Communications, Control and Signal Processing*, March 2008

- Cordina, M.; Debono, C.J. (2009). Maximizing the Lifetime of Wireless Sensor Networks through Intelligent Clustering and Data Reduction Techniques, *Proceedings of the 2009 IEEE Wireless Communications and Networking Conference*, April 2009
- Doherty K.; Adams R.; Davey N. (2007). Unsupervised Learning with Normalised Data and non-Euclidean Norms, *Applied Soft Computing*, vol.7, no.1, pp.203-210, January, 2007
- Ettus M. (1998). System Capacity, Latency and Power Consumption in Multihop-Routed SS-CDMA Wireless Networks, *Proceedings of IEEE Radio and Wireless Configuration*, August 1998
- Fanimokun A; Frolik J. (2003). Effects of Natural Propagation Environments on Wireless Sensor Network Coverage Area, *Proceedings of the 35th IEEE Southeastern Symposium on System Theory*, Morgantown, West Virginia, USA, March 2003
- Handy M.; Haase M. & Timmermann D. (2003). Low Energy Adaptive Clustering Hierarchy with Deterministic Cluster-Head Selection, *Proceedings of the Fourth IEEE Conference on Mobile and Wireless Communications Networks*, Stockholm, September, 2003
- He J.;Tan A.; Tan C. & Sung S. (2003). On Quantitative Evaluation of Clustering Systems, *Information Retrieval and Clustering*, W. Wu, H. Xiong, and S. Shekhar, Eds., Boston, MA, 2003
- Heinzelman W.; Kulik J. & Balakrishnan H. (1999). Adaptive Protocols for Information Dissemination in Wireless Sensor Networks, *Proceedings of the Fifth Annual ACM/IEEE International Conference on Mobile Computing and Networking*, 1999
- Heinzelman W., et al. (2000a). Energy-Scalable Algorithms and Protocols for Wireless Sensor Networks, *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP '00)*, Istanbul, Turkey, June 2000
- Heinzelman, W.; Chandrakasan A. & Balakrishnan H. (2000b). Energy-efficient Communication Protocol for Wireless Sensor Networks, *Proceedings of the Hawaii International Conference System Sciences*, Hawaii, January 2000
- Heinzelman W.; Chandrakasan A. & Balakrishnan H. (2002). An Application Specific Protocol Architecture for Wireless Micro-sensor Networks, *IEEE Trans. Wireless Communications*, vol. 1, no. 4, pp. 660-670, October 2002
- Intanagonwiwat C.; Govindan R. & Estrin D. (2000). Directed Diffusion: A Scalable and Robust Communication Paradigm for Sensor Networks, *Proceedings of ACM MobiCom*, August 2000
- Kohonen, T. (1995). *Self Organising Maps*, Springer-Verlag, Berlin, 1995
- Kotz D.; Newport C.; Gray B.; Liu J.; Yuan Y. & Elliott C. (2004). Experimental Evaluation of Wireless Simulation Assumptions, *Proceedings of the 7th ACM international symposium on Modelling, analysis and simulation of wireless and mobile systems (MSWiM'04)*, Venice, Italy, October 2004
- Maruta T.; Ishibuchi H. (1994). Performance Evaluation of Genetic Algorithms for Flowshop, *Proceedings of the 1st IEEE Conference on Evolutionary computation*, June 1994
- Min R., et al. (2000). An Architecture for a Power Aware Distributed Micro-sensor Node, *Proceedings of the IEEE Workshop on signal processing systems (SIPS'00)*, October 2000
- Molina-Garcia-Pardo J.; Martinez-Sala A.; Bueno-Delgado M.; Egea-Lopez E.; Juan-Llacer L. & Garcia-Haro J. (2005). Channel Model at 868 MHz for Wireless Sensor Networks in Outdoor Scenarios, *International Workshop on Wireless Ad-hoc Networks*, London, UK, May 2005

- Rappaport, T. (1996). *Wireless Communications: Principles & Practice*, Prentice Hall, New Jersey, 1996
- Shepard T. (1996). A Channel Access Scheme for Large Dense Packet Radio Networks, *Proceedings of ACM SIGCOMM 1996*, Stanford University, August 1996
- Shih E., et al. (2001). Physical Layer Driven Protocol and Algorithm Design for Energy-efficient Wireless Sensor Networks, *Proceedings of the 7th Annual ACM/IEEE International Conference on Mobile Computing and Networking*, Rome, Italy, July 2001
- Woo A.; Culler D. (2001). A Transmission Control Scheme for Media Access in Sensor Networks, *Proceedings of the 7th Annual ACM/IEEE International Conference on Mobile Computing and Networking*, Rome, Italy, July 2001
- Xu Y.; Heidemann J. & Estrin D. (2001). Geography-informed Energy Conservation for Ad Hoc Routing, *Proceedings of the ACM/IEEE International Conference on Mobile Computing and Networking*, Rome, Italy, July, 2001
- Ye M.; Li C.; Chen G. & Wu J. (2005). EECS: An Energy Efficient Clustering Scheme in Wireless Sensor Networks, *Proceedings of IEEE International Performance Computing and Communications Conference (IPCCC)*, 2005
- Ye W.; Heidemann J. & Estrin D. (2002). An Energy-Efficient MAC Protocol for Wireless Sensor Networks, *Proceedings of IEEE Infocom 2002*, New York, June 2002