

An Information Management Protocol to Control Routing and Clustering in Sensor Networks

Marzieh Veyseh, Belle Wei and Nader F. Mir

San José State University, SAD

This paper develops and analyzes a novel clustering protocol, the Decentralized Energy Efficient cluster Propagation (DEEP) protocol that manages the communication of data while minimizing energy consumption across sensor networks. The paper also presents an Inter-Cluster Routing protocol (ICR) that is compatible with the proposed clustering technique. The DEEP protocol takes advantage of the multi-rate capabilities of 802.11a, 802.11b, and 802.11g technologies by elevating the data rate to higher levels for shorter transmission ranges. This approach reduces the energy consumption by lowering the transmission time. In order to conserve energy and prolong network lifetime, the DEEP protocol starts with an initial cluster head and gradually forms clusters throughout the network by controlling the dimension of clusters and the geographical distribution of cluster heads. Because this model results in a balanced load among cluster heads, protocol overhead due to frequent re-clustering is eliminated. Simulation results demonstrate that the DEEP protocol distributes energy consumption approximately eight times better than the LEACH protocol-clustering scheme. In addition, the DEEP protocol substantially reduces total data communication and route setup energy consumption in the network compared to the LEACH protocol.

Keywords: information processing for sensors, routing protocol, network management, clustering in sensor network, energy management.

1. Introduction

Advances in semiconductor technologies and wireless communication make the deployment of a network of ubiquitous tiny sensor nodes possible. These self-configuring sensor networks can gather comprehensive information about a large geographical area or about movements of an object for surveillance purposes.

Target tracking, environmental monitoring, system control, and chemical/biological detection are some of the possible applications [12]. A sensor's small physical size implies limited space for a battery. The high density and random distribution of sensors make it impossible to periodically recharge or exchange batteries. As a result, it is critical to use energy-efficient algorithms to prolong the life of sensor networks.

Energy-aware protocols have been developed exclusively for wireless sensor networks [1-11] in medium access control (MAC), routing, and clustering which specifies the network topology. Most of the proposed energy-aware MAC protocols aim to use one of the two methods for reducing communication energy: adjusting the transmission power [6], and keeping the transceiver off as long as possible [11].

Clustering algorithms specify the topology of a hierarchical network that is partitioned into non-overlapping clusters consisting of neighboring nodes. Each cluster has a cluster head; communication among nodes or with a central base station is through this cluster head. Since every node in a clustered network is connected to a cluster head, the route discovery process among cluster heads is sufficient to establish a feasible route in the network. For a large sensor network, clustering can simplify the multi-hop route discovery process and limit the number of transmissions compared to a flat, non-clustered network. The added overhead is the formation and maintenance of clusters.

Clustering techniques can be either centralized or decentralized. Centralized clustering algo-

rithms require each sensor node to send its individual information (such as energy level and geographical position) to the central base station. Based on a predefined algorithm, the base station then calculates the number of clusters, their sizes, and the cluster heads' positions, and informs each node of its newly assigned duty [1], [2]. Given the assumption that sensor networks consist of thousands of nodes, it is impractical, if not impossible, for the base station to collect information about every single node in the network prior to the route setup phase. Therefore, centralized clustering is not an option for large sensor networks.

Decentralized clustering techniques create clusters without the help of any centralized base station. Bandyopadhyay and Coyle [6] present an energy efficient hierarchical clustering algorithm where each sensor node becomes a cluster head with a probability of p and advertises its candidacy to nodes that are no more than k -hops away from the cluster head. Given the limited transmission range of wireless sensor nodes, a hierarchical structure with an arbitrary number of levels has its limitation. As the number of hierarchical levels grows, it is possible that the distance between upper-level cluster heads increases to the point that they are no longer able to communicate with one another. The Max-Min d -Clustering [7] algorithm creates d -hop clusters, but there is no consideration for energy optimization. The Low Energy Adaptive Clustering Hierarchy (LEACH) algorithm [9] is another decentralized clustering algorithm that does not offer a complete energy optimization solution as it has no strategy for specifying cluster heads' position and distribution.

For cluster-based routing protocol, both inter-cluster and intra-cluster routing protocols have been developed to keep the network connected. Assuming that all sensors can act as a relay node, there could be a large number of possible routes from one data source to the sink. To find the proper path, two major approaches can be used:

- Centralized route selection algorithm.
- Distributed route selection algorithm, which is classified into two main categories:
 - Proactive routing
 - Reactive routing

Centralized route selection algorithms aim to choose the appropriate next neighbor for each node using a central command node. The central command node collects the information about direct paths' cost and geographical position of the nodes and finds the least cost path using shortest path algorithms such as Dijkstra [20]. CSEAR is a protocol proposed in [21] that chooses TDMA along with centralized routing to minimize energy consumption. On the other hand, wireless networks that perform distributed routing leave the route selection decision to the sensor nodes by themselves. Nodes with proactive route selection algorithms such as link state routing and distance vector routing [20] keep a routing table that contains next hop information to every single node in the network. Reactive routing protocols set the route to the desirable destination only when it's needed. For Ad-hoc wireless networks, Dynamic Source Routing (DSR) and Ad-Hoc On-Demand Vector Routing (AODV) are the examples of reactive routing [17]. But none of these protocols consider the crucial energy consumption consideration introduced in wireless sensor network.

Another group of on-demand reactive routing protocols have been proposed to address the exclusive issues of wireless sensor network. SPIN [22] and Directed Diffusion [23] introduce a concept of "interest" propagation whenever a node wants to send data or a source needs to ask for it. Flooding the network with interest signals will establish a path from a sink to every possible source (spanning tree). While directed diffusion reinforces the paths with higher data rate, SPIN concentrates exclusively on the path set up via negotiation.

To summarize, both energy conscious clustering and routing algorithms should reduce energy consumption, balance the load among sensor nodes, and prolong network lifetime. This paper proposes a Decentralized Energy-Efficient cluster Propagation (DEEP) protocol that establishes clusters with uniformly distributed cluster heads. This protocol balances the load among all the cluster heads by keeping the clusters' radii fairly equal. This protocol is completely decentralized and there is no need for any location finder device or hardware. This paper also proposes an inter-cluster routing protocol that is compatible with the clustering algorithm that sets the most efficient route among cluster

heads. The remainder of this paper is organized as follows. Section 2 describes the communication energy model used for this paper. Section 3 introduces the DEEP protocol and states its advantages. Section 4 presents an inter-cluster routing protocol. Section 5 evaluates the performance of the protocol and compares it with a general version of the LEACH protocol.

2. Communication Energy Model

2.1. Energy Model for Multi-Rate 802.11

Recent advancements in the area of wireless local area networks have led to the development of today's popular IEEE standards such as 802.11a, b, g. These standards have been widely adopted by consumers and provide the widest possible range of data rates of 54, 48, 36, 24, 18, 12, 9, and 6 Mbps reflecting the trade-off between the transmission range and data rate intrinsic in a wireless communication channel. An accurate energy model is crucial for the development of an energy efficient clustering and routing protocols. The popular energy model of ωd^n cannot properly demonstrate the energy consumption versus transmission distance for wireless communication. In general, all transceiver energy components are summarized as:

$$E_{(Watt)} = \theta + \eta_{amp} \times \omega d^n \quad (1)$$

where θ is a distance-independent term that accounts for the overhead of the radio electronics and digital processing and ω models the free space path loss [15]. Based on environmental conditions, n could be a number between 2 and 4. η_{amp} represents the amplifier inefficiency factor and specifies the inefficiency of the transmitter when generating ωd^n power at the output of the antenna. The effect that a distance-dependent term has on the total energy consumption depends on the real world transceiver parameters, θ , η_{amp} , and the path attenuation ωd^n . If the value of θ overshadows $\eta_{amp} \times \omega d^n$, then the reduction in the transmission distance through the use of multi-hop communication is not effective [15]. The Atheros 2004 tri-mode chipset [14] is used to monitor the real values for the radio hardware. In theory, the maximum efficiency of a power amplifier

is 48.4%. However, practical implementations show that the power amplifier efficiency is less than 40% [19]. Therefore, θ is calculated using the assumption that $\eta_{amp} = (1/0.4 = 2.5)$. While maximum output power and total power consumption is provided in the manufacturer data sheet, θ could be calculated based on the following formula:

$$\begin{aligned} \theta &= \theta_{TX} + P_{RX} \\ &= (P_{TX} - \eta_{amp} \times \omega d^n) + P_{RX} \quad (2) \end{aligned}$$

Table 1 shows the values for P_{TX} and P_{RX} based on the manufacturer's data sheet and the values for θ and $\eta_{amp} \times \omega d^n$ calculated for the chosen chipset. Despite the fact that the path attenuation energy increases exponentially by the transmission distance, the data illustrates that static power consumption (θ) dominates the path loss and, therefore, causes the total power consumption to remain constant as transmission distance increases.

Mode	Max Output Power, ωd^n (dBm) [14]	Total Power Consumption (W) [14]	θ (Watt)	$\eta_{amp} \times \omega d^n$ (Watt)
802.11a	+14	1.85 (P_{TX}) 1.20 (P_{RX})	2.987	0.0625
802.11b	+21	1.75 (P_{TX}) 1.29 (P_{RX})	2.727	0.3125
802.11g	+14	1.82 (P_{TX}) 1.40 (P_{RX})	3.157	0.0625

Table 1. Energy Consumption Parameters for Atheros2004 tri-mode chipset.

Rate (Mbps)	Max. Range	Rate (Mbps)	Max. Range
1	100 m	18	51 m
2	76.5 m	24	41.25 m
6	64.5 m	36	36 m
9	57 m	48	23.1 m
12	54 m	54	18.75 m

Table 2. Expected data rate of 802.11g technology [13].

On the other hand, one of the current developments in ad hoc wireless communication is enabling devices to operate using different transmission data rates. 802.11a, 802.11b, and

802.11g are some of the wireless networking standards that have multi-rate capabilities [13]. Although sensor nodes in general generate data in low rates, they can transmit the information using wireless high-speed modulation and techniques. Table 2 shows the expected data rate for 802.11g wireless technologies. While exploiting the multi-rate capabilities of wireless standards has never been proposed for sensor networks, this technique can decrease the transmission energy for smaller distances by switching to higher data rates and keeping the transceiver on for a shorter period of time.

$$E_{\text{Joule/bit}} = \frac{(\theta + \eta_{\text{amp}} \times \omega d^m) \text{Joule/sec}}{(\text{Rate})_{\text{bit/sec}}} \quad (3)$$

In this case, energy in terms of Joule/bit reduces discretely as transmission distance shrinks. Figure 1 shows the energy consumption of Atheros2004 radio for 802.11g technology at the constant rate of 1Mbps and the same technology with the multi-rate extension.

As Figure 1 illustrates, due to the large value of θ compared to maximum output power, single rate communication energy consumption remains constant as transmission distance increases while communication energy consumption for multi-rate transmission decreases for shorter transmission ranges. However it does not follow the model of ωd^m . Meanwhile, multi-rate communication would necessitate the presence of a robust rate selection protocol.

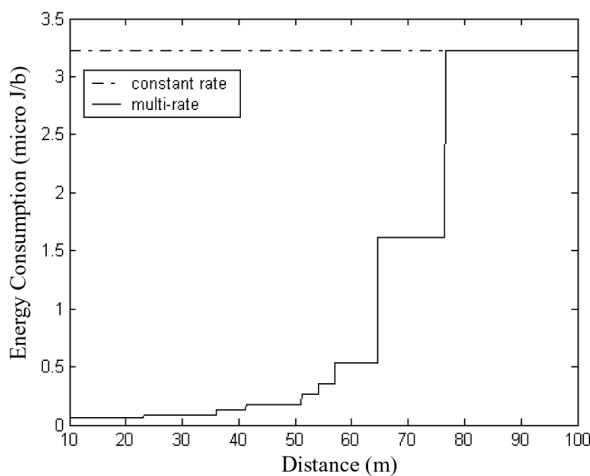


Fig. 1. Energy consumption versus transmission distance for single rate and multi-rate communication (802.11g technology).

2.2. Multihop Communication Efficiency

Considering the real world radio parameters and multi-rate communication impacts, we should re-evaluate the effectiveness of multihop communication. Since multi-rate communication reduces energy consumption for shorter distances by switching to higher data rates, multihop can conserve energy. The traditional objective of multihop communication is to divide the transmission distance into m smaller ranges (m represents number of hops) and relatively conserve energy considering Equation (1) by means of [15]:

$$E = m \times (\theta + \eta_{\text{amp}} \times \omega (d/m)^n) \quad (4)$$

However, if division of transmission distance happens when maximum range is less than 18.75 m for 802.11g, data rate remains constant and total energy consumption multiplies by number of hops. Since sensor networks deal with 2 or even 3-dimensional spaces, the efficiency of multihop heavily depends on the scales and the density of the network. To investigate the efficiency more thoroughly, Figure 2 shows an organization where sensor nodes are placed d meters away from each other, and they tend to send their data packet to the Cluster-Head (CH). We assume that 802.11g technology is being used in this environment. d is an application-dependent parameter and can be chosen based on the sensor's characteristics. It is predicted that for toxin detections, sensors should be placed in average no more than 10 m away from each other. With the choice of 10 m for d , if node B tries to use node A as a relay node and send the data to the

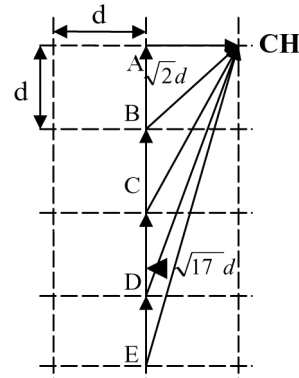


Fig. 2. Sensor nodes' distances from cluster head in a 2-dimensional space.

CH, total energy of the chosen 2-hop path will be larger than the direct transmission energy:

$$\begin{aligned} E(\sqrt{2}d) &= 0.0517 \\ &< E(d) + E(d) = 0.1034 \end{aligned} \quad (5)$$

Also, for nodes C and D, there is no multihop path that can lead to better energy consumption than the direct communication path:

$$\begin{aligned} E(\sqrt{5}d) &= 0.0581 \\ &< E(\sqrt{2}d) + E(d) = 0.1034 \end{aligned} \quad (6)$$

$$\begin{aligned} E(\sqrt{10}d) &= 0.0775 \\ &< E(\sqrt{5}d) + E(d) = 0.1098 \end{aligned} \quad (7)$$

But if node E first sends the data to the intermediate node D, total energy consumption will be less than the direct communication path:

$$\begin{aligned} E(\sqrt{17}d) &= E(41.23) = 0.1789 \\ &> E(\sqrt{10}d) + E(d) = 0.1292 \end{aligned} \quad (8)$$

Node E is placed 41.23 m away from the cluster head. This shows that for nodes that are placed more than 41.23 m away from each other, direct transmission is no more the best possible communication method. In order to compare the energy consumption of direct and multihop communication inside the cluster, we set up an environment representing one cluster. The dimension of the field is $50\text{ m} \times 50\text{ m}$, and 25 nodes

are randomly dispersed in the field. At this point, we assume that cluster head is chosen randomly among the sensors; later details of cluster head selection algorithm will be explained. Figure 3 shows the energy consumption of direct, minimum energy 2-hop, and minimum energy 3-hop path based on the distance between nodes in the cluster and the cluster head. For 802.11g technology, direct transmission is the optimum choice for ranges less than 37 m, which is almost the same as the result from analytical calculations (41 m). However, for ranges greater than 37, minimum energy 2-hop path can lead to significantly lower energy consumption.

3. Proposed Energy-Efficient Clustering Protocol

3.1. Cluster Head Positioning

Based on the results obtained in Section 2.2, for nodes that are placed more than 37 m away from cluster head, a strong path selection algorithm is necessary in order to find the optimum 2-hop or 3-hop path. While direct transmission can eliminate the overhead caused by route setup control messages, its efficiency is limited in terms of transmission range. In order to avoid the frequent control signal transmission and extra power consumption associated with that cluster head could be placed at the center of the cluster, while sensor nodes are positioned closer than

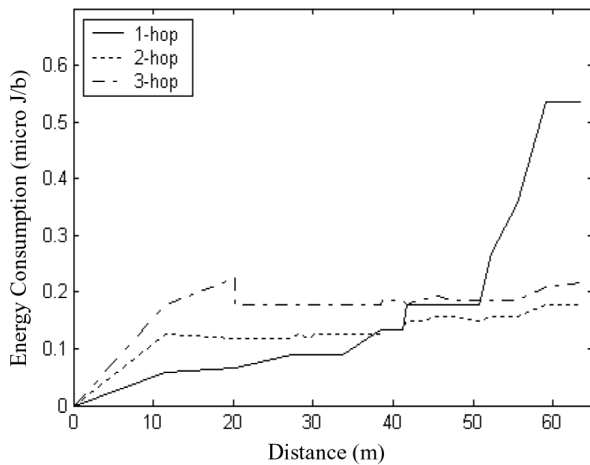


Fig. 3. Communication energy versus distance from cluster head for 802.11g technology.

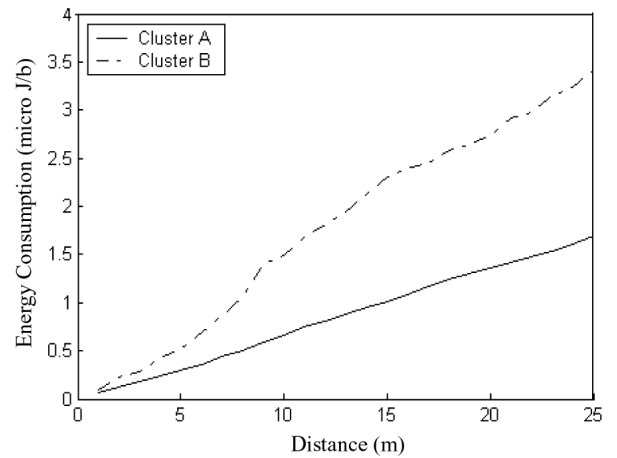


Fig. 4. Cumulative communication energy between sensor nodes and the cluster head versus distance from the cluster head for both centered and cornered positioned cluster heads.

37 m around it. In this case, cluster members can send the data packets directly to the cluster head without the need for any route set up protocol, while efficiency is already achieved through the choice of cluster shape and cluster size. Figure 4 compares two different clusters of the same dimension. *A* is a cluster with the cluster head being positioned at the center, and *B* is a cluster with the chosen cluster head in the corner. The cumulative energy consumption that the centered positioned cluster head should spend to receive one bit of information from all the members is significantly lower than the cumulative energy in the cluster head positioned in the corner.

3.2. Protocol Details

A robust clustering technique is essential in order to configure clusters with almost the same radius and cluster heads that are positioned in the center of clusters. Since sensor nodes begin without any knowledge about their location relative to the base station, distributed clustering algorithm should be able to form clusters without the help of base station and knowledge of network member's position. Although location finder devices and protocols could be deployed, they are either costly or put too much overhead on the network. DEEP is based on the idea of starting with an initial cluster head and forming new cluster head candidates gradually by controlling relative distance between a pair of cluster heads and circular radius of each cluster. Due to the balanced load among cluster heads (will be shown in Section 5.1), periodic re-clustering is not necessary and therefore operational expenses caused by frequent re-clustering is eliminated. In order to explain the details of this algorithm, first we introduce control signals and protocol parameters:

Control signals:

- Cluster Head (CH) declaration signal
- Cluster Head (CH) exploration signal
- Membership search signal

Control Parameters:

- Declaration range (d_r)
- Exploration range (e_{r1}, e_{r2})
- Minimum number of members (m_n)

- E_{rc1} and E_{rc2}

Protocol control parameters are application specific choices and can be defined prior to the network deployment. Protocol DEEP forms clusters by starting with an initial cluster head that can be chosen prior to the network deployment. This initial cluster head starts the cluster set up phase by propagating cluster head declaration signal within the range of d_r . This means that cluster head candidate chooses the appropriate data rate and signal output power so that it can reach nodes that are less than d_r away from the sender. At this point sensor nodes that receive the declaration signal accept the corresponding cluster head as a leader. They can estimate their relative distance to the candidate by looking at the received signal energy level. Once they know the relative distance to the cluster head, they can conserve energy by adjusting the transmission speed to the appropriate value and switching to the sleep mode. Now initial cluster head candidate propagates the cluster head exploration signal within the range of e_{r2} (Figure 5). All the sensor nodes that are in this range can listen to the exploration signal, but only nodes that have never played the role of a cluster head and can verify the following inequality will be chosen as a new candidate:

$$E_r > E_{rc2} \quad \& \quad E_r < E_{rc1} \quad (9)$$

E_r is the received signal energy and it should be bigger than E_{rc1} and smaller than E_{rc2} so that the node can consider itself as a candidate. This ensures that new cluster head candidates are positioned between e_{r1} (m) and e_{r2} (m) away from the initial cluster head. E_{rc1} and E_{rc2} are fixed protocol parameters that can be pre-calculated

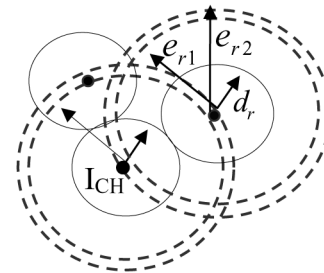


Fig. 5. Initial cluster head starts the advertisement process. New cluster head candidates send the exploration signal within the range of e_{r2} to continue the process of cluster establishment.

and stored in the sensor node memory using the following formula:

$$E_{rc1} = P_{out} - \omega(e_{r1})^n \quad (10)$$

$$E_{rc2} = P_{out} - \omega(e_{r2})^n \quad (11)$$

P_{out} is the constant output power of the cluster head exploration signal, and ω and n are parameters that can be determined based on the environmental conditions of the deployment area. After new cluster head candidates are assigned, each candidate sends a declaration signal within the range of d_r to find new cluster members. If two candidates could hear the declaration signal of each other, they are too close to each other to be considered both as cluster head candidates. Therefore one of them will be eliminated through negotiation, i.e. whenever a cluster head receives a declaration signal, it informs the sender of the message using an acknowledgement message. The cluster head that receives the acknowledgment sends a dissolution message and informs all nodes within the range of d_r about its elimination. If a node receives declaration signal from more than one candidate, it chooses the candidates whose associated signal is received with higher power. Then, confirmed cluster heads propagate exploration signal and search for new cluster head candidates. Nodes that have already been chosen as cluster head or member ignore the cluster head exploration or declaration signals. Therefore, this advertisement process will terminate automatically when all the nodes in the field belong to a cluster. At this point, algorithm might have produced some clusters with very small number of members. Therefore, if there is a cluster that its total number of members is smaller than “Minimum number of members” (m_n), cluster will be dissolved and all the members including cluster head initiate a *Membership search signal*.

Afterwards, they listen to the responds from the local cluster heads and choose the closest cluster head based on the received signal power. At the end, if $t = time_{out}$ and a sensor node hasn't received any control signal, it sends a *membership search signal* and chooses the closest cluster head as a leader. Figure 6 shows the pseudo-code for the DEEP protocol. Protocol execution can be summarized as follows:

1. Initial cluster head finds cluster members by sending “cluster head declaration”.

2. Initial cluster head finds new cluster head candidates by sending “cluster head exploration signal”.
3. Cluster head candidates that are placed on the (e_{r1}, e_{r2}) ring find cluster members.
4. Nodes that receive more than one cluster head declaration choose the closest cluster head based on the received signal energy.
5. Cluster head candidates that receive cluster head declaration signal negotiate with the sender, and one of them gets eliminated.
6. Confirmed cluster heads send “cluster head exploration” to find new cluster head candidates (go to 4).
7. If the number of members in a cluster is less than m_n , all the members find new clusters by sending *membership search signal*.
8. At the end, if a node has not received any control signal, it sends *membership search signal*.

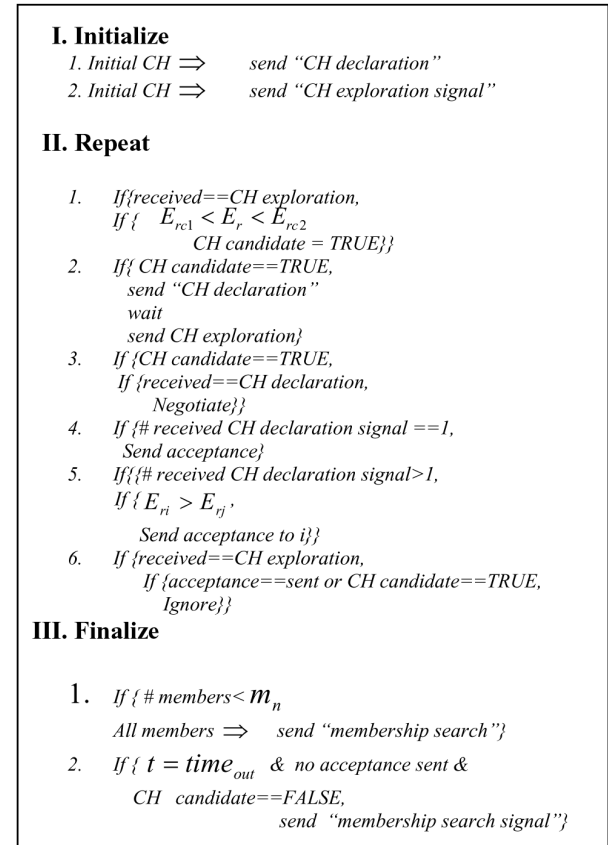


Fig. 6. DEEP protocol pseudo-code.

The protocol provided in this paper has the following advantages:

- As it was discussed, DEEP is a completely decentralized protocol. A sensor node can either elect itself as a cluster head by receiving a cluster head exploration signal or join a cluster by receiving a cluster head declaration signal.
- After the execution of DEEP, all the sensor nodes are covered and belong to only one cluster.
- For the execution of DEEP, there is no need for any location finder hardware such as GPS (Global Positioning System), or a position estimation protocol that puts extra overhead on the sensor nodes.
- DEEP can control the cluster head distribution across the sensor network through the protocol execution methodologies. These tactics include: cluster head candidates should receive the cluster head exploration signal with the certain amount of energy, and if they can hear the declaration signal of each other, one of the candidates will be eliminated.
- Communication cost is minimized via proper selection of protocol parameters (declaration range, exploration range, and minimum number of members).
- Intra-cluster communication is controlled by cluster heads, and nodes transmit their data directly to the cluster heads. Therefore, there is no additional control signal exchange associated with route selection and maintenance inside the cluster.
- Due to the uniform distribution of cluster heads, communication cost of a direct transmission path between a pair of neighboring cluster heads is almost identical across the sensor field. This is one of the most important characteristics of this protocol that contributes tremendously to the ease of deployment of the inter-cluster routing protocol.

4. Inter-Cluster Routing Protocol

After establishing well-distributed cluster heads and clusters in the network, energy conscious

routing is essential in order to set communication routes among cluster heads in a two level hierarchical system. Due to the limited transmission range associated with low power wireless technologies, cluster heads' data packets cannot reach to the base station unless other cluster heads act as relay nodes. In this section, we propose an Inter-Cluster Energy Conscious Routing (ICR) protocol that is compatible with the proposed clustering protocol. ICR uses interest flooding similar to directed diffusion [23] and Energy Aware Routing (EAR) [18] to establish routes between the base station and sensor nodes, but it differs from EAR and directed diffusion in some aspects. First we describe the ICR scheme and then accentuate the differences from EAR.

4.1. Routing Protocol Details

The protocol ICR is a destination-initiated reactive routing protocol. This means that a destination, which we call Local Base Station (LBS), initiates an explicit route discovery phase. This initiation includes propagation of an *interest* signal that floods throughout the network and establishes energy efficient routes for the sake of data communication. Based on the application, which can be either periodic data collection or event driven, the *Interest* signal can include the *type* and the *period* of the desired data. For application that necessitates information from specific locations, interest signal also includes the position of the required information.

If the LBS requires some periodic data collection phases, it sets the *period* in which nodes send that specific *type* of information. Monitoring and surveillance application are examples for data collection paradigm. If the LBS requires sensor nodes to detect one specific event, it includes the *type* of the event in the interest signal (Figure 7). Following the route discovery phase, sensor nodes switch to the sleep mode and wait for the specific event. In case of event detection, non-cluster head nodes send the data directly to the associated cluster head,

Type	Period	Sender's address	Cost field
------	--------	------------------	------------

Fig. 7. Interest signal structure.

and that cluster head uses the previously established route to send the information back to the LBS. We can explain the ICR execution in the following steps:

4.2. Route Discovery Phase

1. The LBS initiates the route discovery by sending an interest signal within the range of R_i . R_i should be high enough to keep the cluster head network connected and low enough to prevent unnecessary energy consumption and interest generation. Due to the even distribution of cluster heads achieved by clustering protocol, R_i can be chosen slightly bigger than the average distance between a pair of adjacent cluster heads. The LBS should adjust the output power and data rate of the interest signal to limit its transmission range to R_i (As discussed in Section 2.) Also, the cost field is set to zero before interest propagation starts.

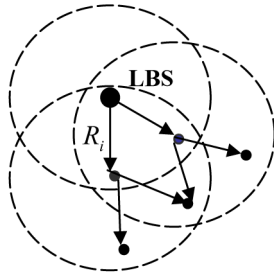


Fig. 8. Local Base Station starts the route discovery process by generating interest signals.

2. Since the distance between a pair of neighboring cluster heads is approximately the same across the network, communication energy consumption associated with two distinct adjacent cluster heads is also the same. Therefore, the cost or weight of a multi-hop path is defined exclusively by the number of hops. In addition, the remaining energy in the cluster heads along the path affects the route selection decision. The total cost function is defined as

$$Cost = \alpha \cdot Hop_Number + \beta \cdot \sum_i \frac{B_M}{B_{ri}} \quad (12)$$

In this equation, B_r represents the remaining energy in the battery of the node, B_M shows

the maximum battery capacity of the sensor nodes, and α and β are normalization factors. Second part of the cost field favors the paths that include nodes with higher energy. To update the cost field, each intermediate cluster head calculates the inverse of its remaining battery power plus one (increment in the number of hops) and adds the outcome to the existing cost value.

3. Each intermediate cluster head that receives the interest signal saves the interest in its memory including the address of the nodes that has sent the message. Then the node should update the cost field of the outgoing interest signal and send it within the range of R_i . All the cluster heads that are placed within this range around the sender can hear the incoming signal.
4. If a cluster head receives an interest that currently exists in the memory i.e. the type, and period is the same, however the senders address is different, it compares the cost field of the received signal with the cost field of the previously saved message. If the incoming interest signal includes a cost field smaller than the previously saved message, the node replaces the old interest entry, updates the cost field, and propagates the packet since the new signal represents a shorter or more energy efficient path. If the new interest signal represents a path with higher number of hops, the node should destroy the packet. In case these two values are equal, the node creates another entry for the incoming interest and keeps that as an option for data communication.

4.3. Data Acquisition phase

1. After each cluster head collects the requested information from sensor nodes and compresses them into a packet with fixed length, it searches for the relay neighbor's address in the memory and sends the packet to that neighbor.
2. In order to reduce the diffusion of spare data bits in the network, relay nodes can receive the data packets each of length L , from N nodes and aggregate them in to one single packet of length L . This will reduce the

number of data bits forwarded by the relay node from NL to L . To enable data aggregation during the data collection period, cluster heads that are closer to the base station, i.e. cost field of the saved interest message includes less number of hops, should wait for their neighbors to send their data packets, then compress the incoming information with their own data and send the packet with the fixed length to the relay neighbor.

This protocol is different from the proposed protocol in [18] (EAR) in two aspects:

1. In the EAR protocol, sensor nodes save and propagate most of the incoming interest signals and only eliminate the ones with very high cost field. However, in ICR, every time that the cost field of the incoming interest message is higher than the previously saved one, the packet gets destroyed. This puts a limit on the generation of the interest messages.
2. In EAR, in order to ensure that the optimal path does not get depleted and network degrades evenly, multiple paths are found between a source and a destination. Each node has to wait for all the interest signals to come and then calculate the average cost between oneself and the destination. Based on the average cost, each path is assigned a probability of being chosen. Depending on the probability, each time one of the paths is chosen. ICR assumes that data aggregation is executed among cluster heads and no packet moves along the chosen path independently. This means that during the data collection period each cluster head aggregates the data from its N adjacent cluster heads and has to forward only one compressed packet rather than N distinct packet. After the execution of routing protocol a spanning tree is established that is rooted in the base station and connects all the cluster heads to the base station. Therefore only the least cost path or the optimum path is a final established route for each cluster head. This way the degradation of the optimal path for each packet is prevented.

5. Simulation Results

We implemented the proposed DEEP protocol along with the ICR routing to evaluate their performance in four aspects:

- Cluster head constellation and distribution of the load across the network
- Re-clustering operation and its overhead
- Optimum density of cluster heads in terms of inter-cluster and intra-cluster communication energy
- Performance comparison between the DEEP protocol and LEACH after the execution of the routing protocol

5.1. Cluster Head Constellation and Distribution of the Load

In this scenario, 3000 sensor nodes are randomly distributed in a field of $550\text{ m} \times 550\text{ m}$, therefore density of the sensors is about 1 per $10\text{ m} \times 10\text{ m}$ area, which is the maximum detection range for the toxin sensors. Considering the results shown in [18], MAC layer abstraction would not degrade the accuracy of simulation results if designed MAC protocol assigns a unique channel for every node and prevents possible collisions. With this assumption, we extracted the MAC layer from our simulations and data packets were sent directly from network layer of one node to the network layer of the neighbor. We simulated the algorithm using different protocol parameters d_r , e_{r1} , e_{r2} and m while initial CH is placed at the center of the field.

Figure 9 shows the output of one of these simulations with parameters $d_r = 30\text{ m}$, $e_{r2} = 80\text{ m}$, $e_{r1} = 78\text{ m}$, $m = 14$. Based on the results obtained from Section 2.2, 30 m is an initial choice for d_r .

In order to avoid overlapping between clusters, the value of e_{r1} and e_{r2} should be more than twice of the value of d_r . Since average distance between sensor nodes in this application is 10 m , 80 m is a fair choice for e_{r2} . The width of the (e_{r1}, e_{r2}) ring should be large enough to accommodate new cluster head candidates and small enough to avoid cluster head candidates that are too close to each other. We chose 2 m

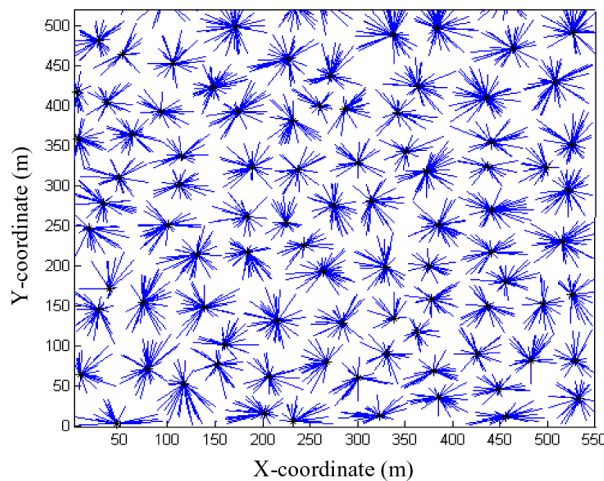


Fig. 9. Initial cluster head, which is placed in the center of the sensor field, starts the simulation by advertising its candidacy. Cluster heads are gradually dispersed across the network. This picture shows the final generated clusters while nodes are directly connected to the associated cluster heads.

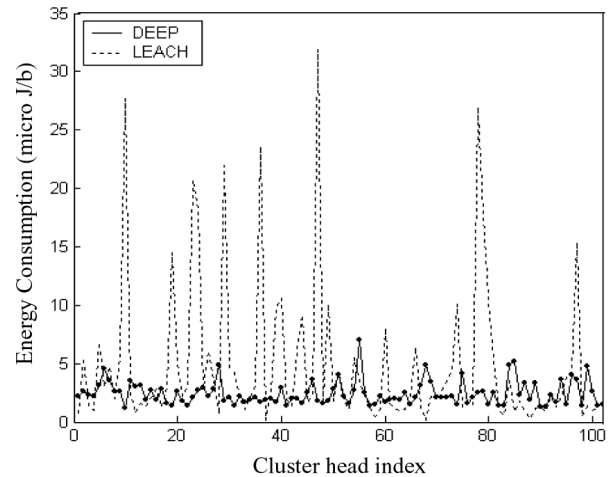


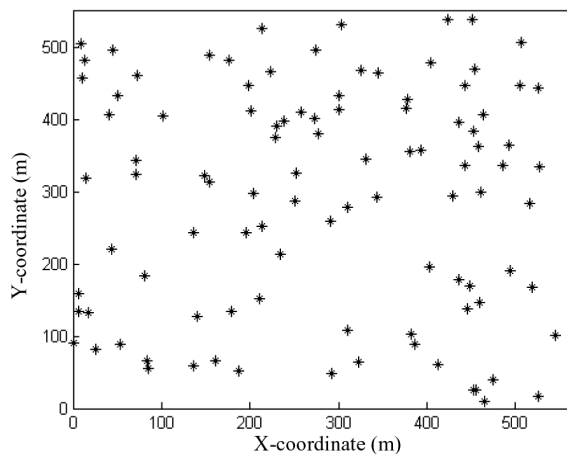
Fig. 10. Energy (μJ) consumed by each cluster head to receive one bit of information from all of the sensor nodes in the cluster.

as an initial value for the ring width. In order to balance the load among cluster heads, DEEP controls the cluster head distribution, rather than the number of cluster members. Although cluster heads that manage more number of members should execute more signal processing for the sake of data aggregation, digital processing consumes much less energy than wireless transmission and there won't be any over-utilized cluster head using this protocol.

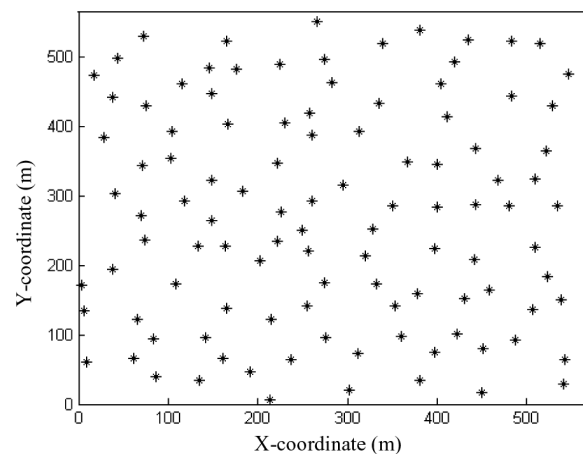
Figure 11 demonstrates the cluster head distribution achieved using protocols LEACH and DEEP. Because of the random selection of cluster

heads in LEACH, some of the cluster heads are too close to each other while others are way too far. This type of cluster head selection causes a lot of burden on some cluster heads and drains their battery very fast.

Figure 10 shows the energy consumption in each cluster head in order to receive one bit of information from all cluster members using 802.11g technology. Due to the inconsistent distribution of cluster heads (Fig. 11), LEACH puts a lot of pressure on some of the cluster heads (CH) while DEEP share out the weight among all of them.



(a) LEACH cluster head constellation



(b) DEEP cluster head constellation

Fig. 11. Comparison between the two protocols, DEEP generates well-distributed cluster heads across the network.

The standard deviation (ESD) among intra-cluster communication energy consumption associated with cluster heads, for LEACH is $7.7013 \mu\text{Joule/bit}$, about 8 times larger than ESD among communication energy consumption for DEEP while the average energy consumption per cluster head for LEACH is $5.476 \mu\text{J/b}$, about 2.3 times larger than the same parameter for DEEP. This can prove the fact that DEEP can perform well in terms of balancing the intra-cluster load among cluster heads compared to the existing protocol LEACH.

5.2. Re-clustering

In order to prevent over-utilization of some sensor nodes, clustering technique should ensure that the cluster head responsibility rotates among all sensor nodes. To achieve this, in LEACH [5], re-clustering is performed periodically, but every round of re-clustering requires several control signal exchanges among self-elected cluster heads and sensor nodes. Re-clustering process in DEEP is based on one small shift in the initial cluster head. When the current period of cluster setting is finished, current initial CH chooses the nearest node that has never been acted as an initial cluster head. This newly chosen initial cluster head starts the clustering process and creates a totally different cluster head constellation. We simulated the clustering process for 30 rounds. Figure 12 shows the accumulative percentage of the nodes that have become a cluster head at the

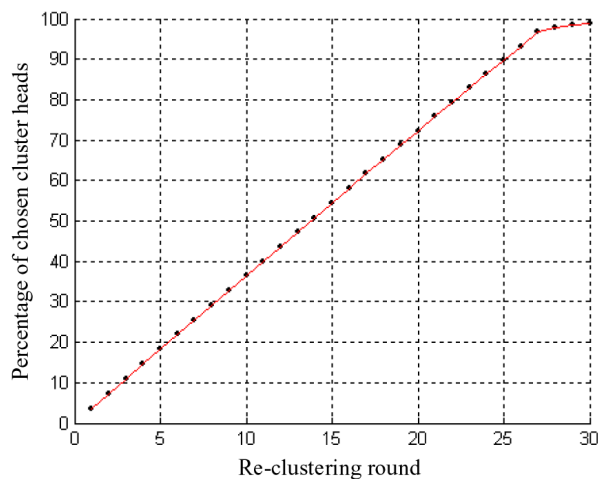


Fig. 12. Accumulative percentage of the nodes that have been chosen as cluster head till the end of each round.

end of each re-clustering round. No node becomes cluster head more than once during the network lifetime. With 3000 sensor nodes, if algorithm generates about 105 cluster heads at each round, after $3000/105 \cong 28$ rounds all the sensor nodes should have become a cluster head. Figure 12 shows that after 27 rounds of re-clustering, number of generated cluster heads will dramatically drop since most of the eligible nodes have been chosen as cluster head and have lost their energy. For this scenario, DEEP can minimize energy consumption associated with re-clustering overheads by reducing the number of necessary rounds to 27 compared to hundreds or even thousands rounds of re-clustering associated with LEACH protocol.

5.3. Optimum Percentage of Cluster Heads

In order to find out the optimum cluster head density and compare the performance of the routing protocol on both DEEP and LEACH, we used a 1600-node network nodes were randomly distributed in a field with area of 400×400 m. In this scenario, sensor nodes send the information directly to their associated cluster head. Each cluster head compresses the data and waits for neighbor cluster heads' data packet. Then, the cluster head compresses all the received data packets into a packet with fixed length and sends it to the relay neighbor. The relay neighbor address has been saved in the node memory through the propagation of interest signal. In-network data aggregation performed by cluster heads as described above helps to reduce the amount of data dispersed in the network.

Our first experiment compares the ESD among energy consumption in cluster heads regarding intra-cluster communication for protocols DEEP and LEACH while percentage of cluster heads in the network is changing. For the sake of fair comparison, we used the communication energy model described in Section 2 for both protocols. Fig. 13(b) approves the fact that DEEP can substantially improve the distribution of the load among cluster heads, regardless of the cluster head density. It also shows that increase in cluster head density can improve the intra-cluster load distribution for both protocols.

Next experiment compares intra-cluster and inter-cluster communication energy associated with

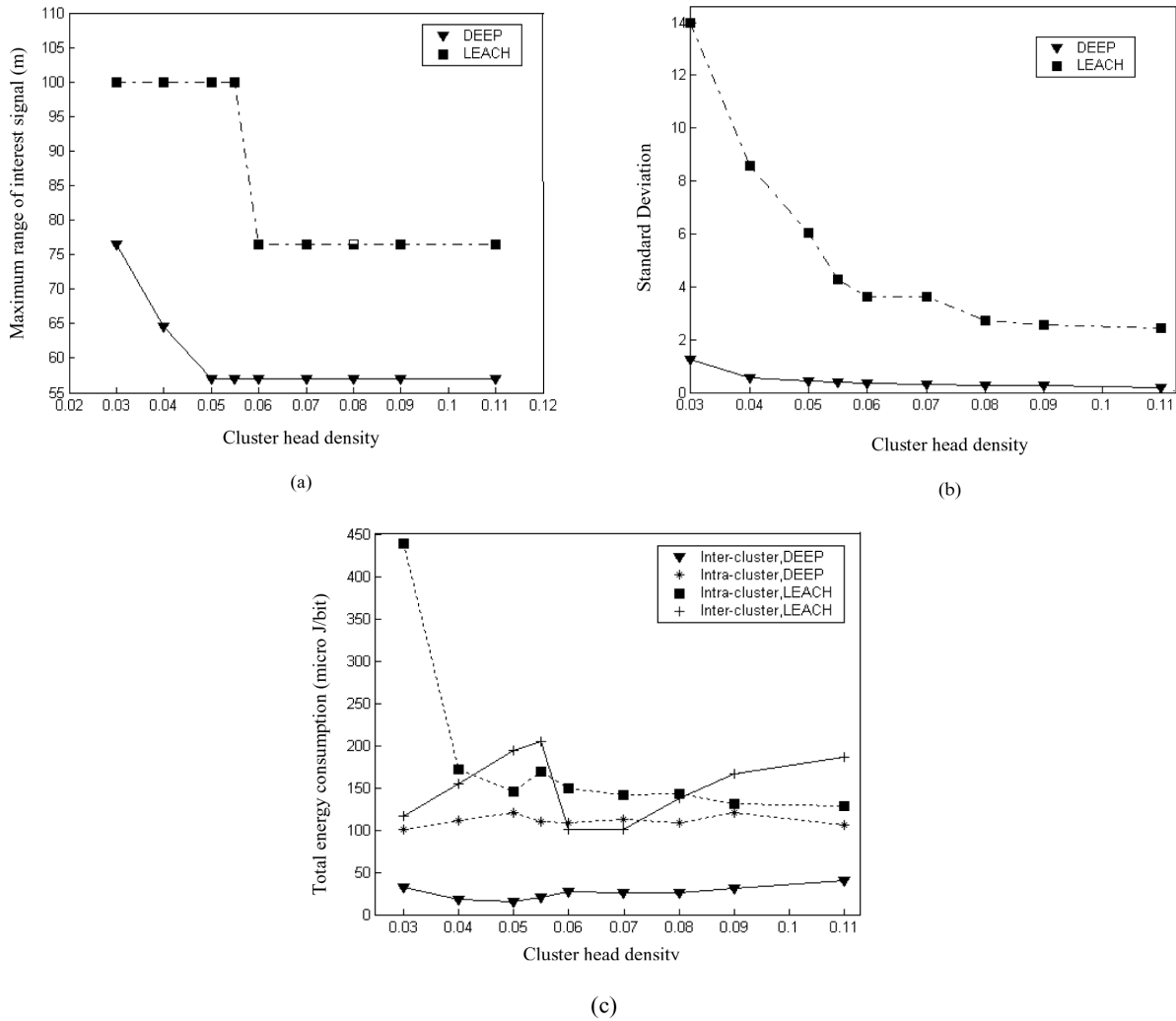


Fig. 13. (a): Maximum range of interest signals; (b): Standard Deviation among intra-cluster energy consumption associated with cluster heads; (c): Total energy consumption inside clusters and among cluster heads.

both protocols. Figure 13(a) shows the maximum range of the interest signal (R_i) that establishes the routes among cluster heads. Since LEACH protocol distributes the cluster heads randomly across the network (Fig. 11(a)), maximum range of R_i should be chosen high enough to make sure that all the cluster heads are covered and network does not get partitioned. But as density of cluster heads increases and they get closer to each other, R_i can decrease. Since we adjust the maximum range of the output signal by changing the transmission data rate, R_i maximum value should follow the values provided in Table 2 for 802.11g. Therefore, as Figure 13(a) shows, R_i changes from 100 m to 76.5 m, for LEACH, when cluster head density reaches 0.055 point and from 76.5 m to 57 m, for DEEP, when cluster head density passes 0.03 point. For

a chosen cluster head density, R_i is higher for LEACH compared to DEEP because DEEP protocol distributes the cluster heads consistently.

Higher values for R_i , in LEACH, lead to extra route setup overheads and faster network degradation. Figure 13(c) shows the total inter-cluster communication energy consumption compared to total intra-cluster energy consumption. The figure shows the reduction in inter-cluster energy consumption when cluster head density varies from 0.03 to 0.05 for DEEP protocol. We can observe the cause of this effect in Fig. 13(a). Since maximum range of interest signal (R_i) reduces from 76.5 m to 57 m, and cluster heads get closer to each other, energy consumption reduces. We can observe the same phenomena for LEACH protocol from point 0.05 to point 0.06.

While R_i is remaining constant, inter-cluster energy increases with cluster head density, because larger number of cluster head indicates larger number of communications. We can observe the increase in inter-cluster energy consumption when cluster head density varies from 0.06 to 0.11. For lower cluster head percentage LEACH has significant consumption compared to DEEP because cluster heads get further away from each other and both inter-cluster and intra-cluster transmission energies increase. To investigate the optimum density of cluster heads in terms of energy consumption, Figure 14(a) shows the total energy consumption in the network to collect one bit of data from every sensor node.

Also Figure 14(a) shows the total overhead caused in the network by interest flooding and route setup for a network that has been clustered using DEEP and LEACH. Also we can observe the significant improvement in the network energy consumption using DEEP protocol. Considering the results from Figure 13(c) and inverse growth of inter-cluster and intra-cluster energy consumption for DEEP protocol, we can see that the summation of these two energies is almost constant while cluster head density is changing. Therefore, the optimum point for cluster head density can be determined based on the route setup energy and distribution of the load. Total flooding overhead associated with LEACH is considerably higher than DEEP, because of the random distribution of cluster

heads and mandatory high value for maximum transmission range of interest signals. Due to the reduction in the maximum range of interest signals (Figure 13(c)), route setup energy drops from 805 to 410 μJ , for LEACH protocol, when density passes the 0.055 point. At the same time, flooding overhead associated with DEEP drops from 89.6 to 40.1 μJ when density reaches the 0.05 point. Route setup energy increases as cluster head density grows more than 0.05 due to the generation of interest signals. From this experiment, the optimum point of operation is around 0.05 for cluster head density.

However, another important parameter to consider is distribution of load. Figure 14(b) shows the standard deviation among energy consumption in cluster heads regarding inter-cluster and intra-cluster communication along with the standard deviation among energy consumption in cluster heads regarding route setup communication. From this figure, we can observe that optimum point for route setup and inter-cluster energy consumption is achieved when cluster head density is about 0.05. However, the standard deviation among energy consumption regarding intra-cluster communications declines as cluster head density increases. Since this reduction does not lead to a better optimum point, we conclude that 0.05 is the optimum point for cluster head density in terms of both load distribution and total energy consumption.

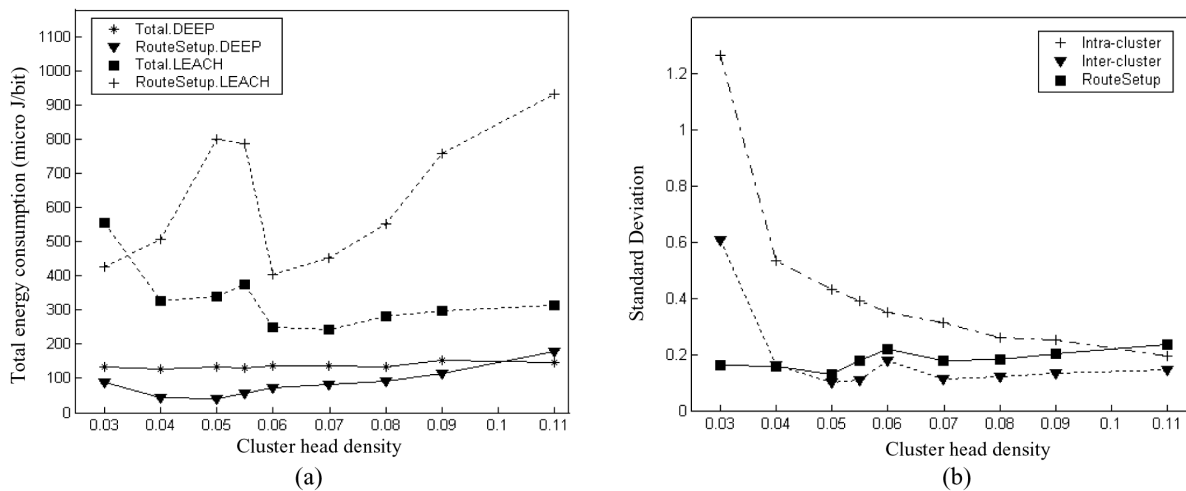


Fig. 14. (a): Total data communication energy consumption in the network along with interest flooding total energy consumption; (b): Standard Deviation among cluster heads' intra-cluster, inter-cluster and route setup energy consumption per cluster head.

6. Conclusion

While wireless sensor network introduces challenging energy limitations, clustering protocols can reduce the amount of energy consumption by dividing the network into well-distributed clusters. In this work, we proposed a new energy-efficient clustering protocol (DEEP) that is based on the idea of controlling the geographical dimensions of clusters and distribution of cluster heads. Because of the balanced load among cluster heads, there is no need for frequent re-clustering, but after current cluster heads are out of energy, protocol can rotate the cluster head position among all the sensor nodes. Also, identical distance between a pair of neighboring cluster heads leads to the ease of route setup deployment. After establishing the routes among cluster heads using Inter-Cluster energy conscious Routing (ICR) protocol, results from our experiments show that DEEP can reduce the energy consumption and distribute the load as nearly as 8 times better than an existing clustering protocol LEACH. Optimum point of operation, in terms of cluster head density, is about 0.05.

Another IEEE standard, 802.15.4, has also developed in order to address issues associated with low-rate personal area networks. The new standard, 802.15.4 allows less complex, inexpensive solutions to be implemented for a wide range of devices. What we are proposing as a new cluster-based routing protocol is a design methodology that can be extended to other standards such as 802.15.4. However, 802.15.4 offers low data rate communication and leads to higher energy consumption during the data transmission phase. At the same time, due to the low complexity associated with 802.15.4, energy consumption during sleep mode or in terms of J/s is much lower than the same parameter for 802.11 Standard. Therefore, we suggest deployment of 802.11 for applications that require data transmission phases with high transmission frequency and deployment of 802.15.4 for applications that require sensor node to stay in sleep mode most of the time.

References

- [1] HALGAMUGE, M.N., et. al., Energy Efficient Cluster Formation in Wireless Sensor Networks, *IEEE 10th Int. Conf. Telecom*, vol. 2, pp. 1571–1576, March 2003.
- [2] TILLET, J., RAO, R., SAHIN F., Cluster-Head Identification in Ad Hoc Sensor Networks Using Particle Swarm Optimization, *IEEE Int. Conf. on Personal Wireless Communications*, pp. 201–205, Dec. 2002.
- [3] GUPTA, G., YOUNIS, M., Fault-Tolerant Clustering of Wireless Sensor Networks, *IEEE Wireless Commun. and Networking Conf.*, vol. 3, pp. 1579–1584, March 2003.
- [4] GUPTA, G., YOUNIS, M., Load-Balanced Clustering of Wireless Sensor Networks, *IEEE Int. Conf. on Comm.*, vol. 3, pp. 1848–1852, May 2003.
- [5] GHIASI, S., et. al., Optimal Energy Aware Clustering in Sensor Networks, *MDPI* 2002.
- [6] BANDYOPADHYAY, S., COYLE, E. J., An Energy Efficient Hierarchical Clustering Algorithm for Wireless Sensor Networks, *Twenty-Second Annual Joint Conf. of the IEEE Computer and Commun. Soc.*, vol. 3, pp. 1713–1723, April 2003.
- [7] AMIS, D., et. al., Max-Min D-Cluster Formation in Wireless Ad Hoc Networks, in *Proceedings of IEEE INFOCOM*, March 2000.
- [8] TAEK, J. K., et. al., Efficient Flooding with Passive Clustering-an Overhead-Free Selective Forward Mechanism for Ad Hoc/Sensor Networks., *Proceedings of the IEEE*, vol. 91, Issue 8, pp. 1210–1220, Aug. 2003.
- [9] RABINER, W., et. al., Energy-Efficient Communication Protocols for Wireless Microsensor Networks, *Hawaii International Conf. on System Sciences (HICSS '00)*, Jan. 2000.
- [10] CHEN, B., et al., Span: An Energy Efficient Coordination Algorithm for Topology Maintenance in Ad Hoc Wireless Networks, in the *Proc. of the 7th ACM Mobile Computing and Comm.* (MobiCom 2001), Rome, Italy, July 2001.
- [11] WEI, Y. J., et al., An Energy-Efficient MAC Protocol for Wireless Sensor Networks, *IEEE Proc. on Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2002)*, vol. 3, pp. 1567–1576, June 2002.
- [12] AKYILDIZ, F., et. al., A Survey on Sensor Networks, *IEEE Communications Magazine*, Vol. 40, Issue. 8, Aug 2002.
- [13] Broadcom white paper, *The New Mainstream Wireless LAN Standard*, softcopy at <http://whitepapers.zdnet.co.uk/0,39025945,60072368p-39000522q,00.htm>
- [14] MEHTA, S., et al., A CMOS Dual-Band Tri-Mode Chipset for IEEE 802.11a/b/g Wireless LAN, *2003 IEEE Radio Frequency Integrated Circuits (RFIC) Symposium*, pp. 427–430, June 2003.

- [15] MIN, R., et. al., Top Five Myths about the Energy Consumption of Wireless Communication, *Mobile Comput. and Commun. Review*, vol. 1, Num. 2, 2003.
- [16] YOUSSEF, M.A., YOUNIS, M.F., ARISHA, K.A., A Constrained Shortest-Path Energy-Aware Routing Algorithm for Wireless Sensor Network, *IEEE Wireless Commun. Networking Conf.*, vol. 2, pp. 794–799, March 2002.
- [17] PERKINS, C.E., et. al., Performance Comparison of Two On-Demand Routing Protocols for Ad Hoc Networks, *IEEE Personal Communications*, vol. 8, Issue 1, pp. 16–28, Feb. 2001.
- [18] SHAH, R.C., et. al., Energy Aware Routing for Low Energy Ad Hoc Sensor Networks, *IEEE Wireless Commun. Networking Conf.*, vol. 1, pp. 350–355, March 2002.
- [19] S. C. CRIPPS, *RF Power Amplifiers for Wireless Communications*, Artech House Publishers, Boston, 1999.
- [20] L. PETERSON, B. S. DAVIE, *Computer Networks*, Morgan Kaufmann, San Francisco, 2003.
- [21] M. A. YOUSSEF, et al., A Constrained Shortest-Path Energy-Aware Routing Algorithm for Wireless Sensor Networks, *IEEE Wireless Commun. Networking Conf.*, 3 (2002), pp. 794–799.
- [22] E. WOODROW, W. HEINZELMAN, SPIN-IT: a data centric routing protocol for image retrieval in wireless networks, *International Conference on Image Processing*, 6 (2002), pp. 913–916.
- [23] C. INTANAGONWIWAT, et. al, Directed diffusion for wireless sensor networking, *IEEE/ACM Transactions on Networking*, 2 (2003), 2–16.

Received: November, 2004

Accepted: January, 2005

Contact address:

Marzieh Veyseh
San José State University
One Washington Square
San José, California
USA

Phone: 95192-0080

e-mail: marziehv@yahoo.com

Belle Wei

San José State University
United States

Nader F. Mir

San José State University
United States

e-mail: nmir@email.sjsu.edu

MARZIEH VEYSEH received her B.Sc. degree in electrical engineering from Tehran University, Iran in 2002. She received her MSc. degree in electrical engineering in 2004 from San José State University in California. She was a research assistant in San José State University, working on the design and development of energy efficient communication protocols for wireless sensor network. Since January 2005 she has been working as a network engineer in Doradus Technologies, San José, California. She has been conducting research in wireless communication networks with an emphasis on sensor networks, computer networks, switching systems and internet protocols. She has been a member of Tau Beta Pi and Phi Kappa Phi since April 2004 and has contributed in editing numerous technical papers for Journal of Computing and Information Technology and for various IEEE conferences.

BELLE W.Y. WEI received an A.B in biophysics in 1977 from the University of California- Berkeley, an M.S. in engineering from Harvard University in 1980, and a Ph.D. in electrical engineering from UC-Berkeley in 1987. She is currently the dean of the College of Engineering at San José State University. Her technical interest areas include sensor network protocols, networking routers/switches, VLSI architectures and algorithms for video compression/decompression, built-in self tests, and computer arithmetics. Prior to her appointment as dean, Dr. Wei served as professor and chair of the Department of Electrical Engineering at San José State. She was a visiting associate professor at Stanford University (1993-94) where she conducted research in developing low-power wireless processors for video decompression and graphics display, and developed error-resilient video decompression algorithms. Dr. Wei has been an active member of the IEEE since 1987 and has served as a program committee member for international conferences. In December 2004, she received a Special Congressional Recognition for Outstanding and Invaluable Service to the Community.

NADER F. MIR received his B.Sc. degree (with honors) in electrical engineering in 1985 and his M.Sc. and Ph.D. degrees, both in electrical engineering, from Washington University in St. Louis, MO, in 1990 and 1994 respectively. He is currently an Associate Professor at the Electrical Engineering Department of San José State University, California. Prior to this position, he was an Assistant Professor of electrical and computer engineering at the University of Kentucky in Lexington. From December 1994 to July 1996, he was a research scientist at the Advanced Telecommunications Institute, Stevens Institute of Technology in New Jersey, working on the design of advanced telecommunication networks. From 1990 to 1994 he was with the Computer and Communications Research Center (CCRC) at Washington University in St. Louis and worked as a research assistant on the design and analysis of gigabit switching systems project. His research interests are: analysis of computer communication networks, design and analysis of switching systems, network design for wireless and telecommunication systems, and applications of digital integrated circuits in computer communications. He is a senior member of the IEEE and has served as conference technical program committee member, steering committee member, and session chair of a number of conferences and has published numerous refereed technical papers, all in the field of communications and networking. He has published a book about video communication engineering and another text-book entitled "Data Networks, Design and Analysis". Dr. Mir is the recipient of several awards including the university teaching recognition award and research excellence award. He is now the Editor of Journal of Computing and Information Technology. He is also the associate editor of IEEE Communication Magazine.
