Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2011, Article ID 984795, 10 pages
doi:10.1155/2011/984795

Research Article

Multi-Hop Data Communication Algorithm for Clustered Wireless Sensor Networks

Dilip Kumar[1] and R. B. Patel[2]

[1] *Academic & Consultancy Services Division, Centre for Development of Advanced Computing (CDAC), Ministry of Communication & Information Technology, Government of India, A-34, Phase-8, Industrial Area, Mohali 160071, India*
[2] *Department of Computer Science & Engineering, Deenbandhu Chhotu Ram University of Science & Technology, Murthal 131039, India*

Correspondence should be addressed to Dilip Kumar, dilip.k78@gmail.com

Received 27 September 2010; Revised 12 January 2011; Accepted 25 February 2011

Recently, continued advances in wireless communication technologies have enabled the deployment of large-scale wireless sensor networks (WSNs). A key concern in the design and development of such WSNs is energy consumption. The hierarchical clustering algorithm is a kind of a technique which is used to reduce energy consumption. It can also increase the scalability, stability, and network lifetime. In some clustering schemes, the communication between a sensor node and its designated cluster head (CH) is assumed to be single-hop. However, multihop communication is often required when the communication range of the sensor nodes is limited or the number of sensor nodes is very large in a network. In this paper, we propose a distributed, randomized, multi-hop clustering algorithm to organize the sensor nodes in a WSN into clusters. The data collected by each sensor node communicate with their respective CHs by using multi-hop communication. The selected CHs collect data from member nodes in their respective clusters, aggregate the data, and send it to a base using multi-hop communication. Simulation results show that proposed algorithm efficiently mitigates the hot spot problem in heterogeneous WSN and achieves much improvement in network lifetime and load balance compared to the existing algorithms.

1. Introduction

Recent advances in microelectromechanical systems- (MEMS-) based technology have motivated the deployment of tiny, low-cost sensor nodes that possess sensing, signal processing, and wireless communication capabilities. These nodes can be deployed at a lower cost than traditional wired networks. A wireless sensor network (WSN) consists of hundreds or thousands of small sensor nodes, which are deployed over a hostile, inhabitable, and harsh environment, possibly for a limited period, with a common, objective, and collaborate to provide distributed sensing, storage, and communication services. These sensor nodes can organize themselves in such a way that they act as front line observation for end users placed at a far distance [1, 2].

In homogeneous WSNs, all the sensor nodes are identical in terms of battery energy and hardware complexity. In clustering, it is evident that the CH nodes will be over-loaded with the long-range communication to the base station (BS) or next cluster head (CH). This means extra processing is necessary for data aggregation which results in the CH nodes expiring before other nodes, although it is desirable to ensure that all the nodes run out of their battery at about the same time, so that very little residual energy is wasted when the system expires. One important way to ensure this is to rotate the role of a CH among over all the sensor nodes as proposed in low-energy adaptive clustering hierarchy (LEACH) [3], power-efficient gathering in sensor information systems (PEGASIS) [4], and hybrid energy-efficient distributed clustering (HEED) [5]. However, these protocols have shown poor performance in heterogeneous environment because the low-energy nodes will die more quickly than the high-energy ones.

On the other hand, in a heterogeneous WSN, two or more different types of nodes with different battery energy, communication capability, and functionality are used. The motivation is that the more complex hardware and the extra battery energy can be embedded in few sensor nodes; thereby, the hardware cost of the rest of the network and communication cost of the sensing nodes can be reduced. Therefore, sensor network systems are gradually moving into heterogeneous designs, incorporating a mixture of different kinds of nodes. Many authors have proposed new routing schemes to address the issues of heterogeneity [6–9].

In this paper, we propose multi-hop data communication algorithm (MDCA) to evaluate the performance of heterogeneous WSNs. Here, we assume that all the sensor nodes of the network are equipped with a different amount of energy, which is a source of heterogeneity. Each sensor node transmits sensing data to the base station (BS) through a cluster head (CH). The CHs are selected periodically by different weighted probability. After the selection of CHs, member nodes communicate with their respective CHs by using multi-hop communication. The CHs collect the data from the member nodes in their respective clusters, aggregate the received data, and send it to the BS using multi-hop communication.

The rest of the paper is organized as follows. Section 2 includes a detailed survey of the related research. Section 3 exhibits the detail of the proposed scheme. Simulation results and its discussion are presented in Sections 4 and 5. Finally, Section 6 concludes the paper.

2. Related Work

The benefits of using clustered heterogeneous WSNs, containing different devices with different capabilities, have been presented recently in the literature.

In a clustered architecture, the sensor network is organized as a number of clusters and each member node belongs to only one cluster. The clustering routing algorithm is firstly presented by the LEACH [3], which guarantees that the energy load is well distributed by dynamically created clusters and the CHs are dynamically selected according to a priori optimal probability. In LEACH, during the setup phase, when clusters are being created, each node decides whether to become a CH for the current round. This decision is based on a predetermined fraction of nodes and the threshold $T(s)$, which is given by the following:

$$T(s) = \begin{cases} \dfrac{p_{\text{opt}}}{1 - p_{\text{opt}} \times \left(r \bmod \left(1/p_{\text{opt}}\right)\right)} & \text{if } s \in G, \\ 0 & \text{otherwise,} \end{cases} \quad (1)$$

where p_{opt} is the predetermined percentage of CHs and r is the count of current round. The G is the set of sensor nodes that have not been CHs in the last $1/p_{\text{opt}}$ rounds. Using this threshold, each node will be a CH at some round within $1/p_{\text{opt}}$ rounds. After $1/p_{\text{opt}}$ rounds, all nodes are once again eligible to become CHs. In this way, the energy concentration on CHs is distributed. LEACH does not consider the residual energy of each node so the nodes that have relatively smaller energy remaining can be the CHs. This makes the network lifetime shortened.

Many routing algorithms have been proposed in the past few years to address the challenges of single-hop and multi-hop communication [3, 10–13]. In [1], the authors have presented taxonomy and general classification of published clustering schemes. In [13], an investigation about cluster size and the number of CHs in the region was achieved when all the devices in a WSN are deployed randomly.

Heterogeneous WSN, supported by recent technological advances in hardware enabled the deployment of tiny, low-power sensors along with limited on-board signal processing and wireless communications capacities. WSNs become increasingly useful in many critical applications [1, 2, 14–17], such as nuclear, biological, and chemical attack detection and protection, home automation, battlefield surveillance, and environmental monitoring. Several energy-efficient protocols designed for heterogeneous networks are based on single-hop clustering techniques, which effectively increased the scalability and lifetime for WSNs [6, 7, 14]. The possibility that more than one type of sensor nodes work together in the WSN is first presented by [3]. However, an in-depth study has not been done by the authors. Stable election protocol (SEP) [9], and energy-efficient heterogeneous clustered (EEHC) scheme [6], in both protocols, CHs are selected based on the weighted probability of each node related to the initial energy. Each member node or CH directly communicates with the CH or BS. Therefore, in large network areas, lots of energy will be consumed.

Many research works have been proposed to deal with nodes' limitation problems; they are related to routing within the sensor networks. In [2, 18], the authors have investigated the existing clustering algorithms. It is essential to improve energy efficiency for WSNs, as the energy supply for sensor nodes is usually extremely limited. Clustering is the most energy-efficient organization for wide application in the past few years, and numerous clustering algorithms have been proposed for energy saving [19–21]. In clustered WSNs, two typical methods to aggregate data after it has been collected from all member nodes before the intercluster communication occurs, another is to aggregate data over each passing hop [18, 22]. In [10, 19], the authors have presented multi-hop routing algorithm for intercluster communication. This algorithm is based on multi-hop routing, which worked on the principle of divide and conquer, and performed well in terms of load balance and being energy efficient as compared to LEACH.

In [21], the authors have proposed energy-efficient hierarchical clustering algorithm (EEHCA) for WSN which improves the performance of LEACH and HEED [5], in terms of network lifetime. EECHA adopts a new method for CH selection, which can avoid the frequent election of CH. In order to improve the performance of the sensor network, the authors have introduced a new concept of backup CHs. Therefore, when nodes finished the communication within their own clusters and the CHs have finished the data aggregation, the head clusters will transmit aggregated data to the BS. In [22], the authors have presented different techniques for the selection of CHs in homogeneous WSNs.

In [23], the authors have studied LEACH scheme and proposed two new schemes (i.e., energy-LEACH and multi-hop LEACH). Entergy-LEACH improves the CH selection method and Multi-hop LEACH (M-LEACH) improves the communication mode from single-hop to Multi-hop between CH and BS. Both the schemes have better performance than LEACH scheme.

In [24], the authors have presented a novel vehicular clustering scheme integrating hierarchical clustering on the basis of classical routing algorithm, which is called multilayer clustering routing algorithm (MLCRA).

Therefore, MLCRA proposed for moving vehicles to mitigate the hot spot problem in WSN and achieves much improvement in network lifetime and load balance compared to the old algorithms which are direct, LEACH, and DCHS. In MLCRA, nontop-level data transfer within the cluster use direct means of communication, and the top-level CHs use multi-hop communications.

3. MDCA: The Proposed Scheme

This section describes the detail of the MDCA to meet the demands of a wide range of heterogeneous applications. However, when we consider a general sensor network that may be deployed over a large region, the energy spent in the power amplifier related to distance may dominate to such an extent that using multi-hop mode may be more energy-efficient than single-hop mode.

Our network model is composed of three types of nodes deployed uniformly in a square region, including normal nodes, advanced nodes, and a few super nodes. The selection probability of each node to become a CH is weighted by the initial energy of a node relative to that of the normal node in the network.

We assume each sensor node transmits sensing data to the BS through a selected CH by using multi-hop communication approach. All the CHs are selected periodically by different weighted probability. Each member nodes communicate with their respective CHs by using multi-hop communication (i.e., intracluster communication). The CHs collect the data from the member nodes in their respective clusters, aggregate the data, and send it to the BS using multi-hop communication (i.e., intercluster communication).

3.1. System Model. We make some assumptions about the sensor nodes and underlying network model, which are as follows. (1) All the sensor nodes are uniformly dispersed within a square field. (2) All the sensor nodes and the BS are stationary after deployment. (3) Multi-hop communication. (4) A WSN consists of heterogeneous nodes in terms of node energy. (5) All the sensor nodes are of equal significance. (6) CHs perform data aggregation. (7) The BS has enough energy in comparison with the other nodes in the network.

3.2. Optimal Clustering. The radio model utilized in MDCA is similar to that of LEACH. The energy consumed by the radio in transmitting L bits data over a distance d is given by the following:

$$E_{Tx}(L,d) = \begin{cases} L \times (E_{\text{elec}} + \epsilon_{\text{fs}} \times d^2), & \text{if } d \le d_0, \\ L \times \left(E_{\text{elec}} + \epsilon_{\text{mp}} \times d^4\right), & \text{if } d \ge d_0, \end{cases} \quad (2)$$

where E_{elec} is the energy dissipated per bit to run the transmitter or the receiver circuit. The parameters ϵ_{fs} and ϵ_{mp} depend on the transmitter amplifier model we use. For simplicity, assume the BS is located inside the field, and that the distance of any node to its CH is $\le d_0$. Thus, the energy dissipated in the CH node during a round is given by the following [6]:

$$E_{\text{ch}} = \left(\frac{n}{k} - 1\right) \times L \times E_{\text{elec}} + \frac{n}{k} \times L \times E_{\text{DA}} + L \times E_{\text{elec}} + L \times \epsilon_{\text{fs}} \times d_{\text{BS}}^2, \quad (3)$$

where k is the number of clusters, E_{DA} is the processing cost of a bit report to the BS, and d_{BS} is the average distance between a CH and the BS. The energy used in a member node is given by the following:

$$E_{\text{nonch}} = L \times E_{\text{elec}} + L \times \epsilon_{\text{fs}} \times d_{\text{CH}}^2, \quad (4)$$

where d_{CH} is the average distance between a cluster member and its CH, which is given by the following:

$$d_{\text{CH}}^2 = \int_0^{x_{\max}} \int_0^{y_{\max}} (x^2 + y^2) \times \rho(x,y)\, dx\, dy = \frac{M^2}{2\pi k}, \quad (5)$$

where $\rho(x, y)$ is the node distribution. By combining (3) and (4), we obtain the total energy dissipated in the WSN which is given by the following:

$$E_t = E_{\text{ch}} + E_{\text{nonch}},$$
$$E_t = L \times (2 \times n \times E_{\text{elec}} + n \times E_{\text{DA}} + \epsilon_{\text{fs}} \times (k \times d_{\text{BS}}^2 + n \times d_{\text{CH}}^2)). \quad (6)$$

By differentiating E_t with respect to k and equating to zero, the optimal number of clusters can be evaluated by the following [9]:

$$k_{\text{opt}} = \sqrt{\frac{n}{2\pi}} \sqrt{\frac{\epsilon_{\text{fs}}}{\epsilon_{\text{mp}}}} \frac{M}{d_{\text{BS}}^2}. \quad (7)$$

If the distance of a significant percentage of nodes to the BS is greater than d_0 then we can obtain the following [9]:

$$d_{\text{BS}}^2 = \int_A (x^2 + y^2) \times \frac{1}{A} = 0.765 \times \frac{M}{2}. \quad (8)$$

By using (7) and (8), we derive the optimal probability of a node to become a CH, p_{opt}, which can be computed by the following:

$$p_{\text{opt}} = \frac{1}{0.765} \times \sqrt{\frac{2}{n\pi}} \times \sqrt{\frac{\epsilon_{\text{fs}}}{\epsilon_{\text{mp}}}}. \quad (9)$$

The optimal probability of a node to become a CH is very important. If the clusters are not constructed in an optimal way, the total energy consumed per round is increased exponentially either when the number of clusters is greater or less than the optimal value.

3.3. Cluster Head Selection Mechanism. The optimal probability of a node to become a CH is a function of spatial density when nodes are uniformly distributed over the network. This clustering is optimal in the sense that energy consumption is well distributed among all sensor nodes and the total energy consumption is minimal. Such optimal clustering highly depends on the energy model that we use. For the purpose of this study, we use similar energy model and analysis as proposed in [3].

Let us assume E_0 is the initial energy of each normal node, m fraction of advanced nodes among normal nodes are equipped with α times more energy than the normal nodes, and m_o proportion of super nodes among advanced nodes are equipped with β times more energy than the normal nodes. Note that a new heterogeneous setting has no affect on the spatial density of the network, so the setting of p_{opt} does not change. On the other hand, due to heterogeneous nodes, the net energy of the network is changed. The new initial energy of each super node is $E_0 \times (1+\beta)$ and each advanced node is $E_0 \times (1+\alpha)$. Hence, the total initial energy of the new heterogeneous network setting is given by the following:

$$E_i = n \times E_o \times (1 + m \times (\alpha - m_o \times (\alpha - \beta))). \tag{10}$$

Hence, the total energy of the system is increased by a factor of $(1+m\times(\alpha - m_o\times(\alpha-\beta)))$. The first improvement to the existing LEACH is to increase the epoch of the sensor network in proportion to the energy increment. In order to optimize the stable region of the system, the new epoch must become equal to $(1/p_{\text{opt}}) \times (1 + m \times (\alpha - m_o \times (\alpha - \beta)))$ because the system has $m \times (\alpha - m_o \times (\alpha - \beta))$ times more energy. If the same threshold is set for super, advanced, and normal nodes with the difference that each normal node in G becomes a CH once every $(1 + m\times(\alpha - m_o\times(\alpha-\beta)))/p_{\text{opt}}$ rounds per epoch, each super node in G becomes a CH$(1+\beta)$ and each advanced node in G becomes a CH$(1+\alpha)$ times every $(1 + m \times (\alpha - m_o \times (\alpha - \beta)))/p_{\text{opt}}$ rounds per epoch, then there is no guarantee that the number of CHs per round per epoch will be $p_{\text{opt}} \times n$. So, the constraint of $p_{\text{opt}} \times n$ CHs per round is violated. Our approach is to assign a weight to the optimal probability p_{opt}.

Let us define p_n, p_a, and p_s are the weighted selection probabilities for normal nodes, advanced nodes, and super nodes. Virtually, there are $(1+m\times(\alpha-m_o\times(\alpha-\beta)))\times n$ nodes with energy equal to the initial energy of a normal node. In order to maintain the minimum energy consumption in each round within an epoch, the average number of CHs per round per epoch must be constant and equal to $p_{\text{opt}} \times n$. In the heterogeneous scenario, the average number of CHs per round per epoch is given by the following:

$$\text{CH}_{\text{average}} = (1 + m \times (\alpha - m_o \times (\alpha - \beta))) \times n \times p_n. \tag{11}$$

The weighed probabilities for normal, advanced, and super nodes are, respectively, as follows:

$$p_n = \frac{p_{\text{opt}}}{1 + m \times (\alpha - m_o \times (\alpha - \beta))}, \tag{12}$$

$$p_a = \frac{p_{\text{opt}}}{1 + m \times (\alpha - m_o \times (\alpha - \beta))} \times (1+\alpha), \tag{13}$$

$$p_s = \frac{p_{\text{opt}}}{1 + m \times (\alpha - m_o \times (\alpha - \beta))} \times (1+\beta). \tag{14}$$

In (1), we replace p_{opt} by the weighted probabilities of normal, advanced, and super nodes to obtain the new thresholds so that it can be used to select the CH for each round. Substitute (12) in (1) and we can find a new threshold for normal nodes which is given by the following:

$$T(s_n) = \begin{cases} \dfrac{p_n}{1 - p_n \times (r \times \operatorname{mod}(1/p_n))} & \text{if } s \in G', \\ 0 & \text{otherwise,} \end{cases} \tag{15}$$

where r is the current round, G' is the set of normal nodes that have not become CHs within the last $1/p_n$ rounds of the epoch, $T(s_n)$ is the threshold applied to a population of $n \times (1-m)$ normal nodes. This guarantees that each normal node will become a CH exactly once every $(1+m\times(\alpha-m_o\times(\alpha-\beta)))/p_{\text{opt}}$ rounds per epoch, and that the average number of CHs that are normal nodes per round per epoch is equal to $n \times (1-m) \times p_n$. Similarly, we can find the thresholds for advanced and super nodes.

3.4. Multi-Hop Communication Mechanism. Single-hop and multi-hop communication are two basic communication patterns which are used in WSNs. It was noticed that in the case of single-hop communication the furthest member nodes or CHs tend to deplete their battery energy faster than other nodes in a network. In other words, in single-hop where data packets are directly transmitted to the CH or the BS without any relay, the nodes located farther away have higher energy burden due to long-range communication, and these nodes may die out first. To overcome this problem, we have used multi-hop communication between member nodes and their respective CH and also between CHs and the BS. We have implemented the same communication approach for both inter-cluster and intra-cluster communication as discussed below.

We use the data transmission network by a directed weighted graph $G = \{V, E\}$, where V is a set of nodes and E is a set of edges. Let us assume v_i and v_j are two nodes in the graph. For the edge $e = (v_i, v_j)$, $w(e) = w_{ij}$, which indicates the weight of e. Here, w_{ij} represents the wasting energy of node v_i which is given in the following equation if node v_i transfers data to node v_j:

$$W_{ij} = \begin{cases} L \times \left(E_{\text{elec}} + \epsilon_{\text{fs}} \times d_{ij}^2\right), & \text{if } d_{ij} \le d_0, \\ L \times \left(E_{\text{elec}} + \epsilon_{\text{mp}} \times d_{ij}^4\right), & \text{if } d_{ij} \ge d_0, \end{cases} \tag{16}$$

where d_{ij} is the distance between v_i and v_j. Similarly, if v_j is the second hop node chosen by another node and v_t is the

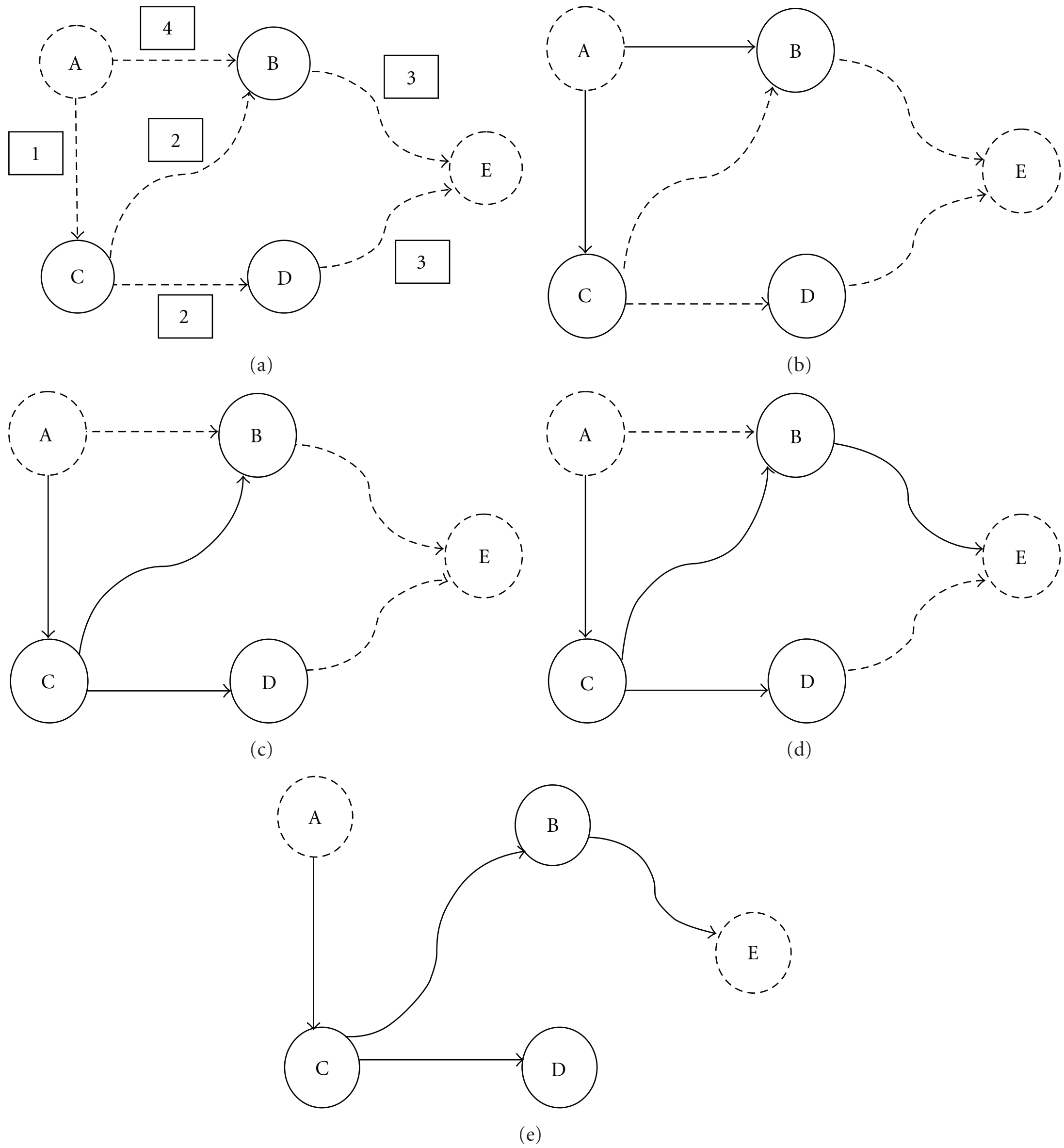

Figure 1: (a) Weighted-directed graph. (b) Shortest path to vertices B, C from A. (c) Shortest path from B, D using C as intermediate vertex. (d) Shortest path to E using B as intermediate vertex. (e) The shortest path calculated between A to E.

BS, then w_{jt} represents the wasting energy of node v_j which is given by the following:

$$W_{jt} = \begin{cases} 2 \times L \times \left(E_{\text{elec}} + \epsilon_{\text{fs}} \times d_{ij}^2\right), & \text{if } d_{ij} \le d_0, \\ 2 \times L \times \left(E_{\text{elec}} + \epsilon_{\text{mp}} \times d_{ij}^4\right), & \text{if } d_{ij} \ge d_0, \end{cases} \quad (17)$$

because it includes a receiving consumption. Therefore, the shortest path weight, also called distance, from v_i and v_t, denoted $d(v_s\ ,\ v_t)$ or d_{st}, is the minimum weight of all possible directed paths with origin v_s and destination v_t.

Let $G = \{V, E\}$ be a directed weighted graph with V a set, whose elements are called vertices or nodes and E is a set of ordered pairs of vertices, called directed edges or arcs. MDCA uses the greedy approach to solve the single source shortest problem. It repeatedly selects from the unselected vertices, vertex v nearest to source s and declares the distance to be the actual shortest distance from s to v. The edges of v are then checked to see if their destination can be reached by v followed by the relevant outgoing edges. The following pseudocode gives a brief description of the working of the MDCA.

Procedure (V: set of vertices 1· · · n {*Vertex 1 is the source*}

Adj[1· · · n] of adjacency lists;

EdgeCost(u, w): edge-cost functions;)

Var: sDist[1· · · n] of path costs from source (vertex 1);

{sDist[j] will be equal to

The length of the shortest path to j }

Begin:

Initialize

{Create a virtual set Frontier to store i where sDist[i]

is already fully solved}

Create empty Priority Queue New Frontier;
sDist[1]←0;{*The distance to the source is zero*}
forall vertices w inV − {1}**do** {*no edges have been explored yet*}
sDist[w] ← ∞
end for;
Fill New Frontier with vertices w in V organized by priorities sDist[w];
end Initialize;
Repeat
v←DeleteMin{New Frontier}; {*v is the new closest; sDist[v]is already correct*}
for all of the neighbors w in Adj[v] **do**
if sDist[w]>sDist[v] +EdgeCost(v,w) then
sDist[w]←sDist[v] +EdgeCost(v,w)
update w in New Frontier {*with new priority sDist[w]*}
end if
end for
until New Frontier is empty
end

The working of MDCA can be explained and understood better using an example. Let us consider that A, B, C, and D are sensor nodes, and E is the destination or the BS as shown in the Figure 1(a). This example will briefly explain each step taken to calculate the shortest distance (sDist).

The above weighted graph has 5 vertices from A–E. The value between the two vertices is known as the edge cost between two vertices. For example, the edge cost between A and C is 1. Using Figure 1(a), we determine the shortest path from the source A to the remaining vertices in the graph.

Initial Step . sDist[A]=0; the value to the source itself
sDist[B]= ∞,sDist[C]= ∞,sDist[D]=∞,sDist[E]=∞;
the nodes not processed yet.

Step 1. Adj[A]={B,C}; computing the value of the adjacent vertices of the graph
sDist[B]=4;
sDist[C]=1;

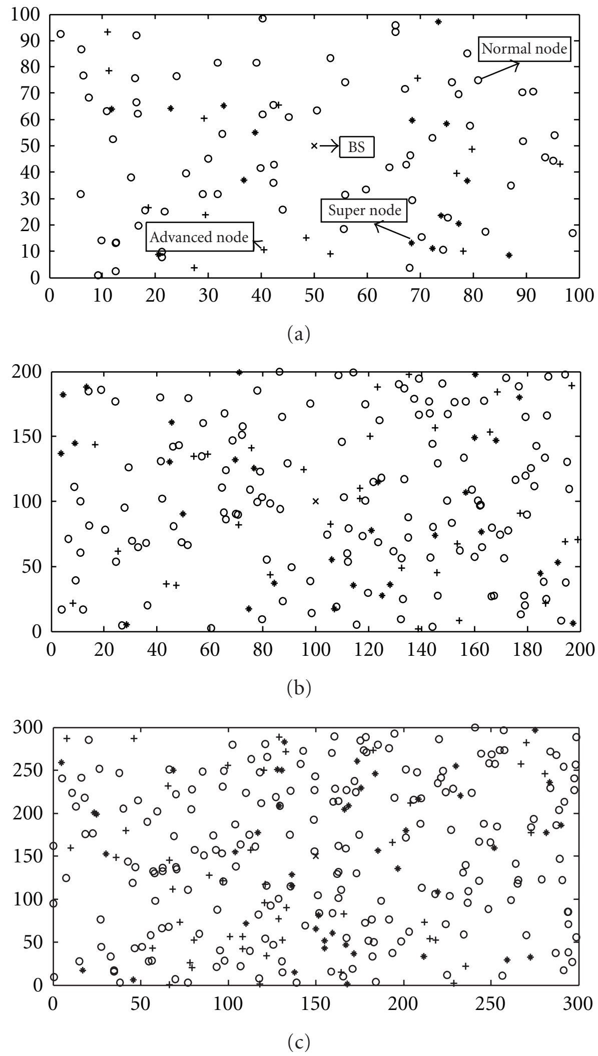

Figure 2: (a) Random deployment of 100 nodes over an area of size 100 × 100 m^2. (b) Random deployment of 200 nodes over an area of size 200 × 200 m^2. (c) Random deployment of 300 nodes over an area of size 300 × 300 m^2.

Step 2. Adj[C]={B, D};
sDist[B]>sDist[C]+EdgeCost[C,B]
4>1+2(True)
Therefore, sDist[B]=3;
sDist[D]=3;
Adj[B]={E};
sDist [E]=sDist[B]+EdgeCost[B,E];
=3+3=6;
Adj[D]={E};
sDist[E]=sDist[D]+EdgeCost[D,E]
=3+3=6
This is same as the initial value that was computed so sDist[E] value is not changed.

Step 3. Adj[E]=0; means there is no outgoing edges from E and no more vertices, algorithm terminated. Hence, the path which follows the algorithm is shown in Figure 1(e).

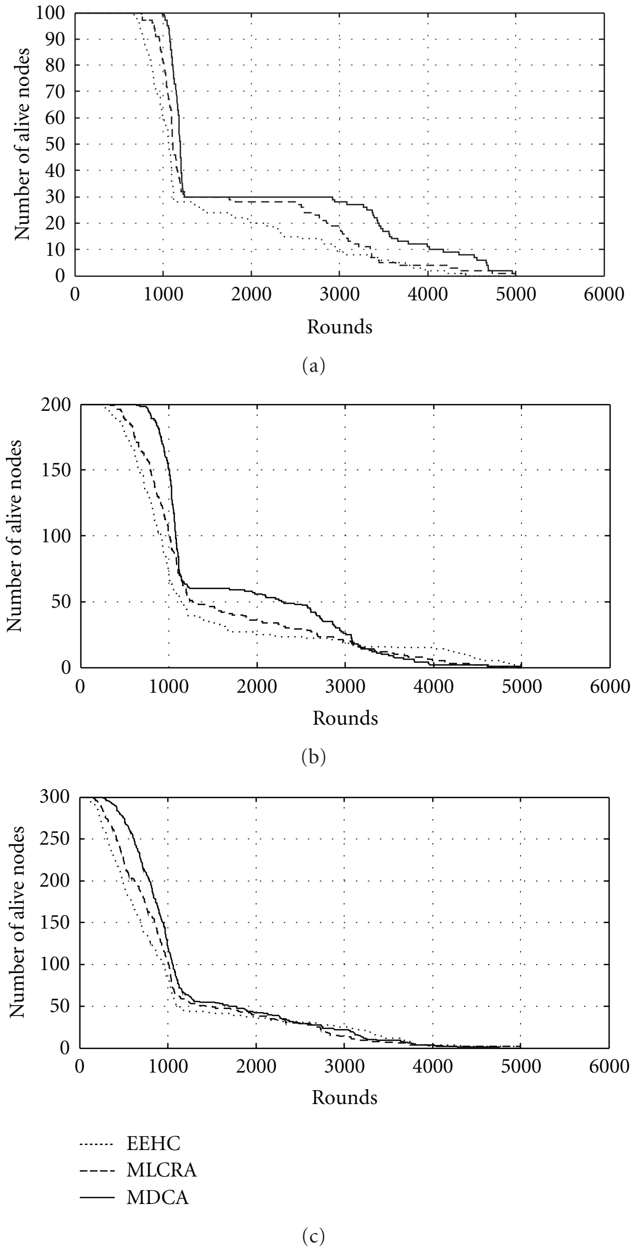

Figure 3: (a) Number of alive nodes per round over an area of size $100 \times 100\,m^2$. (b) Number of alive nodes per round over an area of size $200 \times 200\,m^2$. (c) Number of alive nodes per round over an area of size $300 \times 300\,m^2$.

Finally, the path with the lowest cost link is considered as the shortest-path to the CH or the BS that saves the energy of the network.

3.5. Steady State Phase. Once the clusters are formed, each member node sends data messages in its time slot at the idle state of a frame. In order to avoid collisions during communication, a kind of CSMA model is set up. Instead of transmitting the processed data to the CH directly, every node decides whether to choose another node as the next hop or not. Similarly, each CH decides whether to transmit the data to the BS directly or to send them to the next hop. When a CH has data to send to the BS (i.e., at the end of its frame), it must sense the channel to see if anyone else is transmitting data, if so, the CH waits to transmit the data.

3.6. Traffic Model. The network traffic model depends on the network application and the behaviour of sensed events. The process of data reporting in WSNs is usually classified into three categories: (i) time driven, (ii) event driven and, (iii) query driven. In the time-driven case, sensor nodes transmit their data periodically to the BS. Event-driven networks are used when it is desired to inform the BS about the occurrence of an event. In query-based networks, BS sends a request of data gathering when it is needed. The time-driven scenario is the main focus in MDCA.

4. Simulation Environment

In this section, we evaluate the performance of the MDCA via MATLAB simulations. For simplification, an ideal MAC layer and error-free communication links are assumed [6]. The deployment of the heterogeneous nodes in the network is shown in Figure 2.

4.1. Performance Metrics. The following metrics are used to evaluate the performance of MDCA, EEHC, and MLCRA in different network scenarios.

Network Lifetime. Network lifetime strongly depends on the lifetimes of single nodes that constitute the WSNs. The lifetime of the network basically depends on two major factors: (i) how much energy it consumes over rounds and (ii) how much energy is available for its use. The definition of the network lifetime is determined by the kind of service it provides. In many cases, it is necessary that all the sensor nodes stay alive as long as possible. Since the network performance decreases as soon as a single node dies. In this scenario, it is important to know when the first node dies (FND). Furthermore, sensor nodes can be placed in proximity to each other. Therefore, adjacent nodes could record the same or identical data in the network. Hence, the death of a single or few nodes does not affect the performance of the network. In this case, the metric half node dies (HND) denotes an estimated value for the half-life period of a network. Finally, the metric last node dies (LND) defines network lifetime as the time until all nodes have been drained of their battery energy. This metric is very rarely used in clustering algorithms. Since more than one node is necessary to perform the clustering technique. Hence, in this paper, we use two metrics (i.e., FND and HND) for the evaluation of different algorithms.

Stability. This is the time interval when the first node dies.

5. Simulation Results and Discussion

To validate and compare the performance of MDCA with EEHC and MLCRA, we have conducted simulations for three

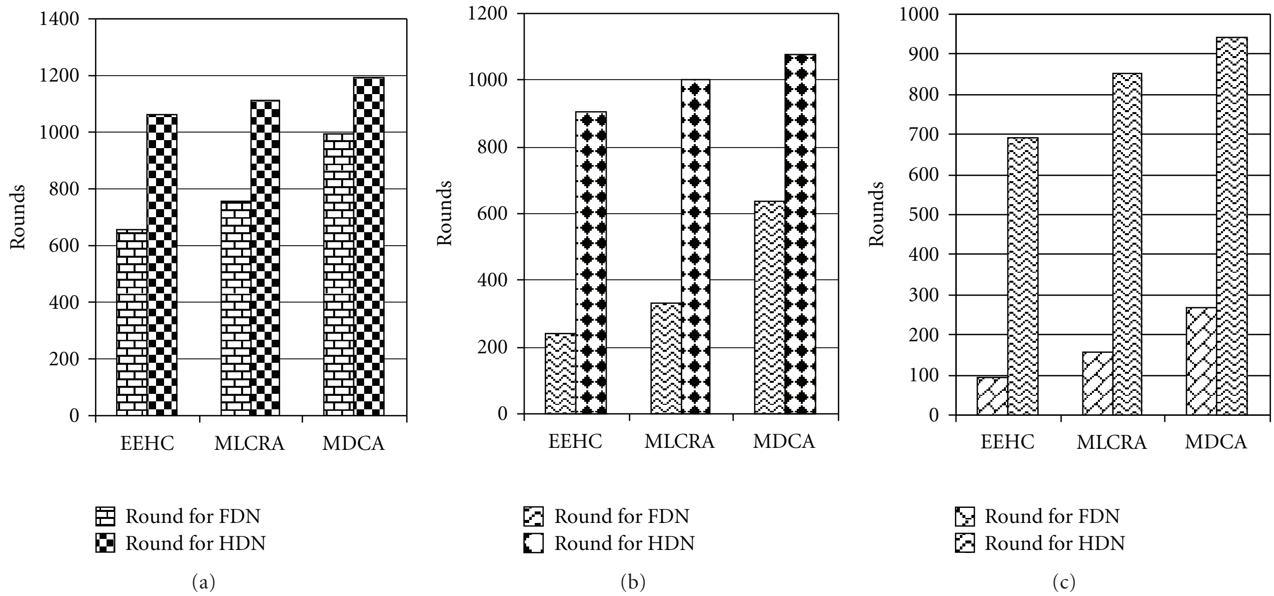

Figure 4: (a) Network lifetime as a function of first dead node and half dead nodes over an area of size 100 × 100 m². (b) Network lifetime as a function of first and half dead nodes over an area of size 200 × 200 m². (c) Network lifetime as a function of first and half dead nodes over an area of size 300 × 300 m².

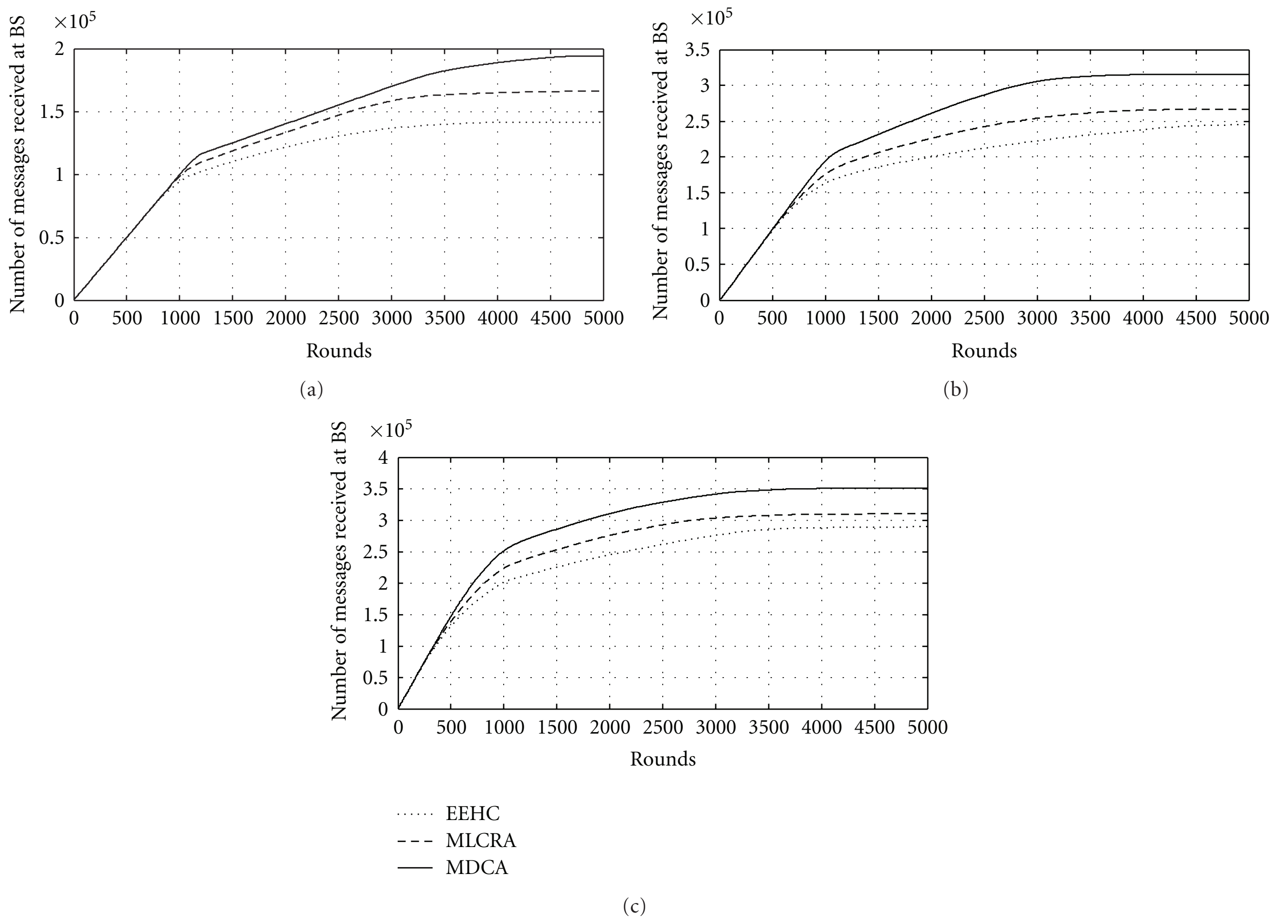

Figure 5: (a) Number of packets received by BS over an area of size 100 × 100 m². (b) Number of packets received by BS over an area of size 200 × 200 m². (c) Number of packets received by BS over an area of size 300 × 300 m².

Table 1: Simulation parameters.

Description	Symbol	Value
The number of nodes	n	100, 200, 300
Proportion of advanced nodes	m	0.3
Proportion of super nodes among advanced nodes	m_o	0.5
Energy factor for advanced nodes	β	2
Energy factor for super nodes	α	3
Initial energy level	E_0	0.25 J
The location of BS	BS	(50, 50), (100, 100), (150, 150)
Data packet Size	L	4000 bits
Network area	$M \times M$	$100 \times 100\,\text{m}^2$, $200 \times 200\,\text{m}^2$, $300 \times 300\,\text{m}^2$
Transmit amplifier if $d_{\text{BS}} \le d_0$	ϵ_{fs}	10 pJ/bit/m^2
Transmit amplifier if $d_{\text{BS}} \ge d_0$	ϵ_{mp}	0.0013 pJ/bit/m^4
Energy dissipated per bit	E_{elec}	50 nJ/bit

scenarios. First, a network with 100 sensor nodes deployed uniformly over an area of size $100 \times 100\,\text{m}^2$, second, a network with 200 sensor nodes deployed uniformly over an area of size $200 \times 200\,\text{m}^2$, and third, a network with 300 sensor nodes deployed uniformly over an area of size $300 \times 300\,\text{m}^2$ as shown in Figure 2. We denote a normal node with "o", an advanced node with "+", a super node with "$*$", and the BS with "$\times$". The simulation parameters are mentioned in Table 1.

Comparing the performance of MDCA with EEHC and MLCRA, when the first node is dead, the merit of MDCA can be clearly seen in Figures 3-4. After a node drains its energy it dies, and it cannot communicate with other nodes any more. We run the simulations for different scenarios and find that MDCA outperforms the existing algorithms. Figure 4 clearly indicates that MDCA can prolong the stability period nearly by 33% for scenario 1, 61% for scenario 2, and 65% for scenario 3 against EEHC and by 23% for scenario 1, 41% for scenario 2, and 46% for scenario 3 against MLCRA, respectively. Since EEHC adopts single-hop communication for both inter-cluster and intra-cluster communication, it shows poor performance in large network areas because all the sensor nodes have to consume more battery energy to perform the long haul communication whereas MLCRA adopts multi-hop communication for inter cluster only.

In EEHC and MLCRA, individual nodes' lifetime is significantly different from each other and hence fails to balance the nodes' lifetime. However, MDCA shows the best performance in balancing the nodes' lifetime since all the sensor nodes die with a closer period with multi-hop communication approach. It is also observed that MDCA shows better performance when the network area increases as indicated in Figure 3. Figure 4 presents that MDCA extend the lifetime of the network over EEHC by 10%, 15%, and 26% and MLCRA by 6%, 8%, and 10%, respectively, for different scenarios when we consider HND as a network lifetime metric.

Figure 5 illustrates the number of messages received by the BS during the network lifetime. The amount of data messages received at the BS will increase over number of rounds as compared with EEHC and MLCRA, respectively. According to the CH selection procedure of MDCA, the network owns uniform number of CHs in every round even after the death of the first alive node whereas in the case of EEHC and MLCRA the number of CHs selection per round become unstable after the death of the first alive node.

6. Conclusion

In this paper, we addressed the problem of developing an energy-efficient multi-hop data communication algorithm (MDCA) for clustered heterogeneous wireless sensor networks (WSNs). The multi-hop communication approach is adopted for both intra-cluster and inter-cluster communication mechanism. Adopting this approach, a path with the lowest cost link is considered as the shortest path either between member nodes and the CH or between CH and the BS to save the energy of the network. Finally, simulation results indicate that MDCA can greatly balance energy consumption of an entire network and thus extends the network lifetime and stability over EEHC and MLCRA, respectively.

References

[1] A. A. Abbasi and M. Younis, "A survey on clustering algorithms for wireless sensor networks," *Computer Communications*, vol. 30, no. 14-15, pp. 2826–2841, 2007.

[2] G. Anastasi, M. Conti, M. D. Francesco, and A. Passarella, "Energy conservation in wireless sensor networks: a survey," *Ad Hoc Networks*, vol. 7, no. 3, pp. 537–568, 2009.

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

[4] S. Lindsey and C. S. Raghavendra, "PEGASIS: power-efficient gathering in sensor information systems," in *Proceedings of the IEEE Aerospace Conference*, vol. 3, pp. 125–1130, 2002.

[5] O. Younis and S. Fahmy, "HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks,"

IEEE Transactions on Mobile Computing, vol. 3, no. 4, pp. 366–379, 2004.

[6] D. Kumar, T. C. Aseri, and R. B. Patel, "EEHC: energy efficient heterogeneous clustered scheme for wireless sensor networks," *Computer Communications*, vol. 32, no. 4, pp. 662–667, 2009.

[7] V. Mhatre and C. Rosenberg, "Homogeneous vs heterogeneous clustered sensor networks: a comparative study," in *Proceedings of the IEEE International Conference on Communications (ICC '04)*, vol. 27, pp. 3646–3651, June 2004.

[8] V. Mhatre and C. Rosenberg, "Design guidelines for wireless sensor networks: communication, clustering and aggregation," *Ad Hoc Networks*, vol. 2, no. 1, pp. 45–63, 2004.

[9] G. Smaragdakis, I. Matta, and A. Bestavros, "SEP: a stable election protocol for clustered heterogeneous wireless sensor networks," in *Proceedings of the 2nd International Workshop on Sensor and Actor Network Protocols and Applications (SANPA '04)*, pp. 251–261, Boston, Mass, USA, 2004.

[10] N. Israr and I. U. Awan, "Multilayer cluster based energy efficient routing protocol for wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 4, no. 2, pp. 176–193, 2008.

[11] N. Khalid and P. B. Alejandro, "ACDMCP: an adaptive and completely distributed multi-hop clustering protocol for wireless sensor networks," *International Journal of Wireless and Mobile Networks*, vol. 2, no. 3, pp. 18–36, 2010.

[12] J. S. Liu and C. H. R. Lin, "Energy-efficiency clustering protocol in wireless sensor networks," *Ad Hoc Networks*, vol. 3, no. 3, pp. 371–388, 2005.

[13] C. Sevgi and A. Kocyigit, "On determining cluster size of randomly deployed heterogeneous WSNs," *IEEE Communications Letters*, vol. 12, no. 4, pp. 232–234, 2008.

[14] S. Bandyopadhyay and E. J. Coyle, "An energy efficient hierarchical clustering algorithm for wireless sensor networks," in *Proceedings of the 22nd Annual Joint Conference on the IEEE Computer and Communications Societies (INFOCOM '03)*, vol. 3, pp. 1713–1723, San Francisco, Calif, USA, April 2003.

[15] W. P. Chen, J. C. Hou, and L. Sha, "Dynamic clustering for acoustic target tracking in wireless sensor networks," in *Proceedings of the 11th IEEE International Conference on Network Protocols (ICNP '03)*, p. 284, 2003.

[16] G. Chen, C. Li, M. Ye, and J. Wu, "An unequal cluster-based routing scheme in wireless sensor networks," *Wireless Networks*, vol. 15, pp. 193–207, 2009.

[17] J. Enrique, D. Melo, and M. Liu, "Analysis of energy consumption and lifetime of heterogeneous wireless sensor networks," in *Proceedings of IEEE Global Telecommunication Conference (GLOBECOM '02)*, vol. 1, pp. 21–25, Taipei, Taiwan, 2002.

[18] C. Stefano, N. Chilamkurti, and S. Zeadally, "A novel centralized clustering algorithm for energy efficient wireless sensor networks," *International Journal of Autonomous and Adaptive Communications Systems*, vol. 1, no. 2, pp. 242–261, 2008.

[19] I. Nauman and I. Awan, "Multi-hop routing algorithm for inter cluster head communication," in *Proceedings of 22nd UK Performance Engineering Workshop*, pp. 24–31, Bournemouth, UK, 2006.

[20] S. Stanislava and W. B. Heinzelman, "Prolonging the lifetime of wireless sensor networks via unequal clustering," in *Proceedings 5th IEEE International Workshop on Algorithms for Wireless, Mobile, Ad Hoc and Sensor Networks (WMAN '05)*, Denver, Colo, USA, 2005.

[21] G. Xin, W. H. Yang, and B. DeGang, "EEHCA: an energy-efficient hierarchical clustering algorithm for wireless sensor networks," *Information Technology Journal*, vol. 7, no. 2, pp. 245–252, 2008.

[22] S. Soro and W. B. Heinzelman, "Cluster head election techniques for coverage preservation in wireless sensor networks," *Ad Hoc Networks*, vol. 7, no. 5, pp. 955–972, 2009.

[23] F. Xianging and S. Yulin, "Improvement on LEACH protocol of wireless sensor networks," in *Proceeding of International Conference on Sensor Technologies and Applications*, pp. 260–264, 2007.

[24] Y. Liu et al., "Multilayer clustering routing algorithm for wireless vehicular sensor networks," *IET Communication*, vol. 4, no. 7, pp. 810–816, 2010.