Hindawi Publishing Corporation
Journal of Computer Networks and Communications
Volume 2015, Article ID 170138, 11 pages
http://dx.doi.org/10.1155/2015/170138

Research Article

An Energy-Efficient Clustering Routing Algorithm Based on Geographic Position and Residual Energy for Wireless Sensor Network

Ali Jorio,[1] Sanaa El Fkihi,[2] Brahim Elbhiri,[3] and Driss Aboutajdine[1]

[1] *LRIT, Research Unit Associated with the CNRST (URAC29), FSR, Mohammed V University, 1014 Rabat, Morocco*
[2] *ENSIAS, Mohammed V University, 713 Rabat, Morocco*
[3] *EMSI, 10000 Rabat, Morocco*

Correspondence should be addressed to Ali Jorio; jorio.ali@gmail.com

Received 27 June 2014; Revised 27 November 2014; Accepted 11 December 2014

Academic Editor: John Doucette

Recently wireless sensor network (WSN) has become one of the most interesting networking technologies, since it can be deployed without communication infrastructures. A sensor network is composed of a large number of sensor nodes; these nodes are responsible for supervision of the physical phenomenon and transmission of the periodical results to the base station. Therefore, improving the energy efficiency and maximizing the networking lifetime are the major challenges in this kind of networks. To deal with this, a hierarchical clustering scheme, called Location-Energy Spectral Cluster Algorithm (LESCA), is proposed in this paper. LESCA determines automatically the number of clusters in a network. It is based on spectral classification and considers both the residual energy and some properties of nodes. In fact, our approach uses the K-ways algorithm and proposes new features of the network nodes such as average energy, distance to BS, and distance to clusters centers in order to determine the clusters and to elect the cluster's heads of a WSN. The simulation results show that if the clusters are not constructed in an optimal way and/or the number of the clusters is greater or less than the optimal number of clusters, the total consumed energy of the sensor network per round is increased exponentially.

1. Introduction

The recent progress in the field of MEMS (microelectromechanical systems), wireless communications, and highly integrated digital electronics has led to the development of microsensors. Such tiny sensors are of low-cost, of low-power, and multifunctional and communicate freely over short distances [1]. These sensor nodes are responsible for the sensing, data processing, and data delivery to the base station (BS). They should work together to form a wireless sensor network (WSN).

A WSN is composed of a large number of sensor nodes, which are randomly or manually deployed in a given coverage area. These nodes collect local physical information, process them, and send them to a BS called sink. For the public notice of the phenomenon, the BS is connected to the internet (Figure 1). Another important characteristic of a WSN is the ability of its nodes to cooperate. Instead of sending raw data to the node responsible for data fusion, the sensor nodes can use their processing abilities to locally carry out calculations and fusion operations to transmit only the information required [2].

These characteristics of wireless sensors enable them to be used in many areas especially for surveillance and monitoring [3]. Compared with the traditional technique of environment monitoring, WSNs technology is a promising green technology for the future in detecting efficiently the environmental variation. WSNs for environment monitoring consist of a large number of low-cost battery-powered sensor nodes, densely deployed throughout a remote or inaccessible physical space [4].

However, the main challenge in the WSNs is the limited power resources of sensor nodes. It is not practical to recharge the nodes batteries or replace them after complete

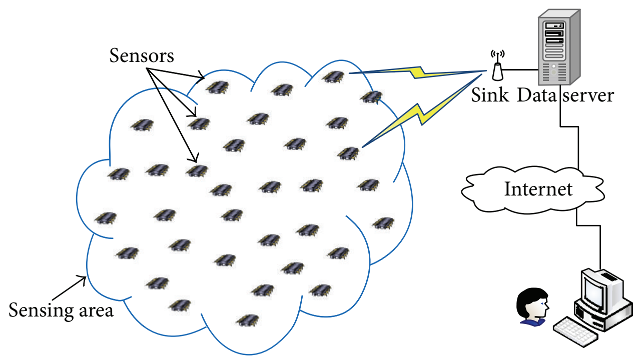

Figure 1: A wireless sensor network model.

depletion of their energy because, in many scenarios, the nodes are deployed in hostile environments [5]. Therefore, while traditional networks aim to achieve a high level quality of service (QoS), the sensor network protocols must focus primarily on energy conservation to maximize the network lifetime [6]. Many research issues have been addressed in this topic. However, design of energy-efficient clustering and routing algorithms is the most promising problem amongst them [7].

Clustering in WSN consists of grouping efficiently the sensor nodes into distinct clusters; each cluster has one leader called cluster head (CH). The CHs collect the data from all their corresponding members, aggregate them, and transmit them to the BS. Each sensor node belongs to one and only one cluster and communicates only with its CH. Therefore, selection of CHs needs to be properly addressed in order to balance the energy consumption of the CHs; otherwise they would die quickly due to the extra work load for data aggregation and data forwarding. Most cluster based routing algorithms select initially the CHs randomly or by probabilities and then form the clusters [7]. However, in such cases, all CHs can be located in a small region of the network and some ordinary nodes will be isolated, which may cause a network dysfunction.

Graph clustering is an area in cluster analysis that looks for groups of related vertices in a graph. Several graph clustering algorithms have been proposed in the last years and spectral clustering algorithms are amongst the most famous. These spectral clustering algorithms have attracted attention over the past few years in many applications, such as image segmentation [8] and social network analysis [9], mainly because of their efficiency and mathematical elegance. Moreover, they have the advantage of providing bounds of graph cut and partitioning problems, due to their spectral relaxation [10]. They are based on Laplacian matrix's eigen-decomposition of either weighted or unweighted graphs [11].

This paper proposes a new energy-efficient environmental monitoring algorithm over cluster based WSNs. The proposed algorithm is called Location-Energy Spectral Clustering Algorithm (LESCA). LESCA is based on spectral clustering by using the graph theory, in order to separate the network to a fixed and an optimal number of clusters. This algorithm consists of three phases: cluster setup, cluster heads selection, and data transmission. First, LESCA selects CHs amongst each cluster using the residual energy, the distance from the BS, and the distance from the cluster's centroid of a sensor node. Then, the cluster's nodes communicate with the selected CH. The latter collects, processes, and transmits the information to the BS. Hence, LESCA aims to extend the network lifetime by distributing energy consumption, minimizing control overhead (to be linear in the number of nodes), and producing well-distributed CHs and compact clusters. Our proposed system considers a denser deployment strategy where the distances between neighboring sensor nodes are quite short. In this way we are aiming to detect environmental variation in a much faster way and send the related information to a BS as quickly as possible.

The rest of the paper is organized as follows. In Section 2, we provide a brief overview of some related research work. Details and properties of the proposed algorithm are given in Section 3 and Section 4 while Section 5 evaluates the performance of LESCA via simulations and compares LESCA with some other clustering protocols. Conclusion and some perspectives are drawn in Section 6.

2. Related Work

Clustering in WSNs is an effective way to minimize the energy consumption of sensor nodes. In fact, clustering performs data aggregation and fusion in order to decrease the number of transmitted messages to the BS and reduce transmission distance of sensor nodes.

Recently, a number of cluster based routing algorithms have been proposed to address the main challenges in WSNs [2, 7, 12]. These algorithms are designed to maintain the information in the network topology, reduce the overhead generated by the discovery of roads, and minimize energy consumption by taking into account the specificity of these networks. They are in most cases energy oriented; that is, they are designed to extend the lifetime of the network, and in some cases they are oriented quality of service (QoS). In this section, we present a brief overview of some related research work.

Low-Energy Adaptive Clustering Hierarchy (LEACH) [13], which is a popular cluster based routing technique, is an energy-efficient communication protocol. In LEACH, the sensor nodes are divided periodically into several clusters. For each cluster, a sensor node is selected as a CH. Thus, LEACH performs a periodic randomized rotation of the CH nodes. The operations of LEACH are generally separated into two phases: the setup phase and the steady-state phase. In the setup phase, CHs are selected and clusters are organized. In the steady-state phase, the data transmissions to the BS take place. The role of the CH is assigned by the node getting a random number between 0 and 1. If the number is less than the threshold values $T(n)$, the node becomes a CH for the current round:

$$T(n) = \begin{cases} \dfrac{P}{1 - P * (r * \bmod (1/P))}: & \text{if } n \in G \\ 0 & \text{otherwise,} \end{cases} \quad (1)$$

where n is the given node, P is the predetermined percentage of CHs ($P = 5\%$), r is the current round, and G is the set of nodes that have not been selected as CHs in the last $1/P$ rounds. Using this threshold, each node will be a CH at some round within $1/P$. After the election of CH nodes, each ordinary node will determine the optimal CH to join in terms of minimum energy required for transmission. However, the random selection of the CH nodes may obtain a poor clustering setup, and the distribution of CH nodes may not be uniform. Thus, some sensor nodes have to transfer data through a longer distance and the reasonable energy saving is not obtained in WSN.

In LEACH-centralized (LEACH-C) [14], which is an enhancement over the LEACH protocol, the BS selects the CH nodes and divides the network into clusters. This is made by collecting the information of the position and energy level from all sensor nodes. Thus, the sensor nodes with energy below the average node energy cannot be CH for the current round. However, like in LEACH, the role of CHs rotates all the time and the optimal location of CH nodes is not guaranteed.

The idea proposed in LEACH has been an inspiration for many hierarchical routing protocols such as EECS [15] and DECSA [16], although some protocols have been independently developed [17–19].

In [17], the authors proposed a cluster scheme based on residual energy and distance from the BS. The idea of this clustering scheme is to modify the k-means algorithm. In cluster formation phase, authors start by randomly selecting k CH nodes. They calculate the distance between each sensor node and the randomly selected CH nodes. Each node is assigned to its near CH that has the highest residual energy. Then, they recompute the CH by using centroid method and repeat the process. Finally, for each cluster, the sensor node, which has the minimum distance from the centroid point, is a new CH. Moreover, the main problem of this scheme is the random selection of the initial CHs. Because of this, it does not provide equitable distribution of energy between clusters.

In addition, the authors of [16] present a distance-energy cluster structure algorithm (DECSA) based on the classic clustering algorithm LEACH. This algorithm considers the distance between the network nodes, the position of the BS, and the residual energy of nodes. Its main idea is to partition the network into three levels of hierarchy. It is made to reduce the energy consumption of CH nodes, resulting from the nonuniform distribution of nodes in the network, and thus avoid direct communications between the BS and CH that has minimal energy and is far away from the BS.

Moreover, in [15], the authors present an energy-efficient clustering scheme (EECS) for periodical data gathering applications in large scale sensor networks. In the cluster head's selection step, several CHs are elected by localized competition, which take into account the residual energy and coverage radius of candidate nodes. Then, in the cluster's formation step, a weighted function for the plain node is introduced to make decision of which proper cluster should be joined. The function is calculated by the cost of intracluster communication and the cost of communication between the CHs and the BS.

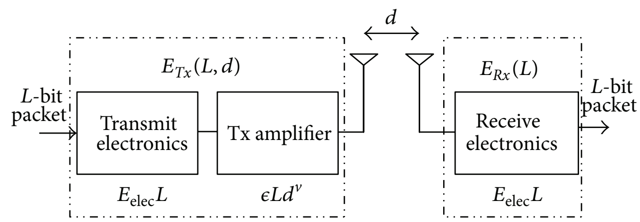

Figure 2: Radio energy dissipation model.

Besides, in [18], the authors propose a partitioning algorithm for WSN based on k-nearest neighbor (KNN). They propose to partition the network into a number of clusters where each cluster contains one BS and all sensor nodes communicate directly with the BS contained within it.

In [20], the authors present an approach called the spectral classification based on Near Optimal Clustering in WSNs. This approach is based on spectral bisection for partitioning a network into two clusters and then applies this method recursively to obtain optimal number of clusters. Spectral bisection method is based on second eigenvalue λ_2 of the Laplacian matrix of the considered graph. Median value of corresponding eigenvector of λ_2 is used to bipart the graph [8]. The authors apply this method recursively to obtain the desired number of clusters. Then, for each cluster, a CH is also elected. This task is assigned in rotational way between nodes without considering their residual energy. However, recursive spectral bisection always produces 2^n clusters, where n is the number of iterations. The approach cannot partition the network into any desired number of clusters.

While there are advantages to use distributed cluster formation algorithms, these protocols offer no guarantee about the position and the number of the CH nodes. However, using a central control algorithm to form the clusters may produce better clusters by dispersing the CHs throughout the network. This is the basis for our proposed algorithm.

3. Radio Energy Dissipation Model and Assumptions

The 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 [13, 14].

According to the radio energy dissipation model illustrated in Figure 2, the amount of energy required for the transmission of L-bit data to a distance d is expressed as follows:

$$E_{Tx}(L,d) = \begin{cases} L \cdot E_{\text{elec}} + L \cdot \epsilon_{fs} \cdot d^2 & \text{if } d < d_o \\ L \cdot E_{\text{elec}} + L \cdot \epsilon_{mp} \cdot d^4 & \text{if } d \geq d_o. \end{cases} \quad (2)$$

When receiving this data, the required energy is rather described as

$$E_{Rx}(L,d) = L \cdot E_{\text{elec}}, \quad (3)$$

where E_{elec} is the energy dissipated per bit to run the transmitter or the receiver circuit, ϵ_{mp} and ϵ_{fs} depend on the used transmitter amplifier model (representing, resp., the multipath fading mode and the free space mode), d_0 is the threshold of both energy models, and d is the distance between the sender and the receiver. By equating the two expressions (2) at $d = d_0$, we obtain

$$d_0 = \sqrt{\frac{\epsilon_{fs}}{\epsilon_{mp}}}. \quad (4)$$

Furthermore, assume area A square meters over which N sensor nodes are uniformly distributed and clustered in K clusters. The total energy dissipated during a round (E_{Round}) in the hierarchical routing approaches is determined by

$$E_{\text{Round}} = \sum_{k=1}^{K} E_{\text{CH}_k} + \sum_{j=1}^{N-K} E_{\text{nonCH}_j}, \quad (5)$$

where E_{CH_k} is the consumed energy when the CH of the cluster k receives, aggregates, and transmits data to the BS. E_{nonCH_j} is the consumed energy by a non-CH node j when this node senses its environment, collects sensed data, and transmits it to the closer CH node.

On the one hand, the E_{CH_k} is defined as follows:

$$E_{\text{CH}_k} = E_{\text{CH to BS}_k} + E_{\text{Reception}_k} + E_{\text{Aggregation}_k}, \quad (6)$$

where

$$E_{\text{CH to BS}_k} = \begin{cases} L \cdot E_{\text{elec}} + L \cdot \epsilon_{fs} \cdot d_{\text{to BS}_k}^2 & \text{if } d_{\text{to BS}_k} < d_o, \\ L \cdot E_{\text{elec}} + L \cdot \epsilon_{mp} \cdot d_{\text{to BS}_k}^4 & \text{if } d_{\text{to BS}_k} \geq d_o, \end{cases}$$

$$E_{\text{Reception}_k} = |S_k| \cdot L \cdot E_{\text{elec}},$$

$$E_{\text{Aggregation}_k} = |S_k| \cdot L \cdot E_{\text{DA}} \quad (7)$$

with

(i) $E_{\text{CH to BS}_k}$ being the energy consumed when the CH of the cluster k transmits data to the BS,

(ii) $E_{\text{Reception}_k}$ being the total energy consumed when the CH of the cluster k receives data from its cluster's nodes,

(iii) $E_{\text{Aggregation}_k}$ being the total energy needed by the CH of the cluster k to process data,

(iv) E_{DA} being the processing cost of a bit report to the BS,

(v) $|S_k|$ being the number of nodes in the cluster k,

(vi) $d_{\text{to BS}_k}$ refers to the distance between the CH of the cluster k and the BS.

If we consider these assumptions, some of the CH nodes operate on the free space mode while the others operate on the amplification mode. Let l be the number of the last category of CH nodes. We consider the following equation:

$$\sum_{k=1}^{K} E_{\text{CH to BS}_k} = L\left(K \cdot E_{\text{elec}} + \epsilon_{mp} \cdot \sum_{i=1}^{l} d_{\text{to BS}_i}^4 + \epsilon_{fs} \cdot \sum_{i=1}^{K-l} d_{\text{to BS}_i}^2\right). \quad (8)$$

Finally, the total energy consumed by all CH nodes is

$$\sum_{k=1}^{K} E_{\text{CH}_k} = L\left((K+N) \cdot E_{\text{elec}} + N \cdot E_{\text{DA}} + \epsilon_{mp} \cdot \sum_{i=1}^{l} d_{\text{to BS}_i}^4 + \epsilon_{fs} \cdot \sum_{i=1}^{K-l} d_{\text{to BS}_i}^2\right). \quad (9)$$

On the other hand, the E_{nonCH_j} is defined by the following equation:

$$E_{\text{nonCH}_j} = \begin{cases} L \cdot E_{\text{elec}} + L \cdot \epsilon_{fs} \cdot d_{\text{to CH}_j}^2 & \text{if } d_{\text{to CH}_j} < d_o \\ L \cdot E_{\text{elec}} + L \cdot \epsilon_{mp} \cdot d_{\text{to CH}_j}^4 & \text{if } d_{\text{to CH}j} \geq d_o. \end{cases} \quad (10)$$

In the network, there are $N - K$ non-CH nodes. Some of these nodes operate on the free space mode while the others operate on the amplification mode. Let m be the number of this category of nodes. We obtain

$$\sum_{j=1}^{N-K} E_{\text{nonCH}_j} = \sum_{j=1}^{N-K-m} L \cdot E_{\text{elec}} + L \cdot \epsilon_{fs} \cdot d_{\text{to CH}_j}^2 + \sum_{j=1}^{m} L \cdot E_{\text{elec}} + L \cdot \epsilon_{mp} \cdot d_{\text{to CH}_j}^4, \quad (11)$$

$$\sum_{j=1}^{N-K} E_{\text{nonCH}_j} = L\left((N-K) \cdot E_{\text{elec}} + \epsilon_{mp} \cdot \sum_{j=1}^{m} d_{\text{to CH}_j}^4 + \epsilon_{fs} \cdot \sum_{j=1}^{N-K-m} d_{\text{to CH}_j}^2\right). \quad (12)$$

From (9) and (12), we conclude that the total energy dissipated during a round in the hierarchical routing approaches is defined by

$$E_{\text{Round}} = L\left(2 \cdot N \cdot E_{\text{elec}} + N \cdot E_{\text{DA}} + \epsilon_{mp}\left(\sum_{i=1}^{l} d_{\text{to BS}_i}^4 + \sum_{j=1}^{m} d_{\text{to CH}_j}^4\right) + \epsilon_{fs}\left(\sum_{i=1}^{K-l} d_{\text{to BS}_i}^2 + \sum_{j=1}^{N-K-m} d_{\text{to CH}_j}^2\right)\right). \quad (13)$$

In a WSN, an important challenge is to reduce the total consumed energy of each round (given by (13)). Thus, if the clusters are not constructed in an optimal way, the total consumed energy of the sensor network per round is increased exponentially, either when the number of clusters that are created is greater or especially when the number of

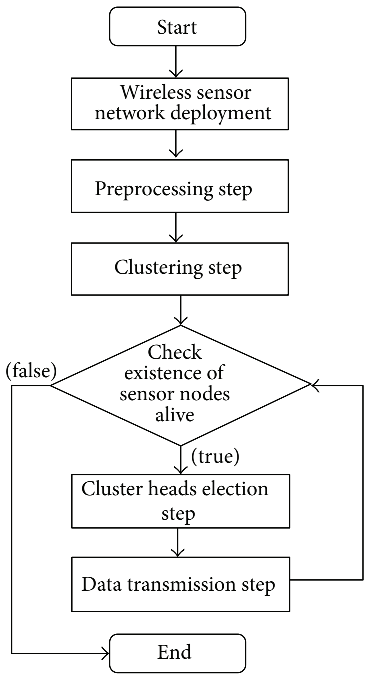

Figure 3: Flowchart of the proposed algorithm.

the constructed clusters is less than the optimal number of clusters. To address these questions, we propose a new protocol using a spectral clustering approach. The next section explains in detail the proposed solution.

4. Proposed Approach: LESCA

Our main objective of the proposed algorithm is to extend the network lifetime through an efficient centralized algorithm. This algorithm guarantees an efficient clustering and CHs selection.

In LESCA, the large number of sensor nodes will be divided into several clusters by using spectral clustering method. The latter has recently gotten great attention in many research areas and has made use of the spectrum of the similarity matrix of the data to cluster a considered set of elements, for example, cluster points using eigenvectors of matrix derived from the data [10]. For each resulting cluster, a sensor node is selected as a CH. The selection of CH nodes is based on a heuristic function. Non-CH nodes can only monitor the environment and send data to its CH. The CH node collects the data from ordinary nodes in the cluster. Then, it processes data and sends them to the BS. An ordinary node cannot send data directly to the BS.

Spectral clustering can be derived as an approximation to such graph partitioning problems. It can be easily applied to partition a WSN. The basic idea of this class of methods is to construct a weighted graph from the data; here the data are sensor network. Its goal is to divide the data points into several homogeneous clusters [11].

In this section, we give details of our proposed LESCA algorithm. The new algorithm consists of four steps: preprocessing step, clustering step, cluster heads election step, and data transmission step (Figure 3). In the remainder, we consider a homogeneous network with N nodes dispersed on a square field to continuously monitor the environment. The BS is located outside the square field.

4.1. The Preprocessing Step. The preprocessing step starts after the sensor nodes are randomly distributed in the sensing area. In the beginning, the BS broadcasts a "Hello" message to all nodes at a certain power level. This way, each node can compute the approximate distance to the BS based on the received signal strength. It helps nodes to select the appropriate power level to communicate with the BS. Then, the latter collects the different nodes positions and triggers the spectral clustering process.

Based on the spectral clustering principle, in the proposed approach, the BS constructs the graph corresponding to the WSN. Indeed, let x be an observation vector composed of the sensor network nodes. This vector can be represented by an undirected graph $G(V, E)$, where V is the set of vertices (sensor nodes) identified by an index $i \in \{1, \ldots, N\}$ and E is the set of edges that link each of the two vertices (communication link). Let $A \in \mathbb{R}^{N\times N}$ be the similarity matrix of the graph G. Each value a_{ij} of A is associated with each pair of the graph nodes (i, j).

This spectral clustering strategy is a relaxation of minimizing Ncut partitioning problem—we want to find a partition of the graph such that the edges between different groups have a very low weight (which means that points in different clusters are dissimilar from each other) and the edges within a group have high weight (which means that points within the same cluster are similar to each other)—and it is based on random walks in the similarity graph. A random walk in a graph is a stochastic process which randomly jumps from one vertex to another. Hence, this spectral clustering strategy can be interpreted as trying to find a partition of the graph such that the random walk stays long within the same cluster and seldom jumps between clusters [21].

The energy consumed for intracluster communications can be extremely saved when those communications are computed by the free space model [14]. Thus, we consider that there is no similarity between two nodes if the distance between them is greater than d_0. The similarity matrix is constructed as follows:

$$A = \left[a_{ij}\right] = \begin{cases} \exp\left(\dfrac{-d^2(i,j)}{2\sigma^2}\right) & \text{if } d(i,j) \le d_0 \\ 0 & \text{otherwise.} \end{cases} \quad (14)$$

The total weight of edges incident to node i is given by $d_{ii} = \sum_{N}^{j=1} a_{ij}$. The degree matrix $D \in \mathbb{R}^{N\times N}$ of G is a diagonal matrix defined by $D = [d_{ij}] = \sum_{N}^{j=1} a_{ij}$. The $N \times N$ Laplacian matrix of the graph is, as introduced by [21], expressed as follows:

$$L = D^{-1} \cdot A. \quad (15)$$

4.2. Clustering Step. The aims of the current step are to determine the optimal number of clusters and to fix the set of the WSN clusters.

The spectral clustering algorithm presented here can be thought of as consisting of three stages [22].

(i) *Preprocessing*. It consists of the normalization of the Laplacian matrix L. We start with a smoothing step to make sure that the matrix is not ill-conditioned.

(ii) *Spectral mapping*. Some eigenvectors of the preprocessed Laplacian matrix are computed. The study of eigenvectors and eigenvalues of a squared matrix is the essence of the spectral theory. Each data point i is mapped to a tuple representing the values of component i in the mentioned eigenvectors. Thus, we construct a new matrix U from the K eigenvectors related to the K largest eigenvalues of the Laplacian matrix.

(iii) *Postprocessing*. It consists of using a grouping algorithm to cluster the data. Hence, we deal with each row of U as a point in $\mathbb{R}^{N\times N}$ and cluster them into K clusters via K-means (that attempts to minimize distortion). The sensor node i is assigned to cluster C_j if and only if row i of the matrix U was assigned to cluster C_j.

Also, the most important question raised by the proposed strategy concerns the optimal number of clusters K that must be considered. For this purpose, we use the spectrum analysis of the Laplacian matrix; it provides global information about the sensor network structure [21]. This spectrum is the graph eigenvectors ordered by the magnitude of their corresponding eigenvalues $\Lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_N\}$. Most stable clustering is given by the value i that maximizes the expression:

$$K = \max \Delta_i = \max \left|\lambda_i - \lambda_{i-1}\right|, \quad \text{where } i = \{1, 2, \ldots, N\}. \tag{16}$$

Hence, the optimal number of clusters K must be

$$K = \max \Delta_i = \left|\lambda_K - \lambda_{K-1}\right|. \tag{17}$$

We notice that in the proposed algorithm we determine the clusters before specifying the CHs. Also, the optimal number of clusters is as well automatically defined. So, our algorithm is completely different from the others. In order to define the number of CHs and cluster's partitioning, the other protocols run the same technique in each iteration and therefore consume more energy.

4.3. Cluster Heads Election Step. Once the clusters are determined, the next step consists in electing the CHs of the network. Here, we consider the total consumed energy in each round (given in (13)). We note that, by considering K cluster, the consumed energy depends on the distances between the CH and the non-CHs of each cluster and the distances between the CHs and the BS. Hence E_{Round} is minimal if the terms in (18) are minimal:

$$\begin{aligned} &\epsilon_{mp}\left(\sum_{i=1}^{l} d^4_{\text{to BS}_i} + \sum_{j=1}^{m} d^4_{\text{to CH}_j}\right) \\ &\quad + \epsilon_{fs}\left(\sum_{i=1}^{K-l} d^2_{\text{to BS}_i} + \sum_{j=1}^{N-K-m} d^2_{\text{to CH}_j}\right). \end{aligned} \tag{18}$$

However, it is known that the amplifier energy in a multipath fading channel model is greater than the amplifier one in a free space model ($\epsilon_{fs} \prec \epsilon_{mp}$). Thus, to minimize E_{Round}, a large part of sensor nodes must operate in a free space model.

Hence, for each cluster, we propose to take into account the position and the energy of the different nodes of the cluster so as to determine its CH. So, the election of the optimal CH of a given cluster can be made in three stages.

Stage 1. Considering the residual energy of each node allows determining the average energy of the cluster E_{cluster}. Besides, a given node can be viewed as a candidate CH if its residual energy is greater than a threshold E_{minCH} (this energy is equal to the energy used to collect, aggregate, and transmit data to the BS) and is given by the following equation:

$$\begin{aligned} E_{\text{minCH}} = {} & \left|S_k\right| * \left(E_{Rx}(L, d) + E_{\text{Aggregation}}\right) \\ & + E_{Tx}(L, d). \end{aligned} \tag{19}$$

So, for each cluster, we select the set of nodes (say S_1) that have a residual energy greater than E_{minCH} and E_{cluster}.

Stage 2. Next, we determine the centroid point of each cluster. These centroid points have new coordinates which are not equal to any position of sensor nodes in the WSN. Thus, for each cluster k, the coordinates of the centroid point are

$$x^k_{\text{centroid}} = \frac{\sum_{i=1}^{|S_k|} x_i}{\left|S_k\right|}, \qquad y^k_{\text{centroid}} = \frac{\sum_{i=1}^{|S_k|} y_i}{\left|S_k\right|}. \tag{20}$$

Then, we select a subset of the candidate CHs from S_1 that are close to the centroid point. So as to take into account the error that could be caused by the system, we introduce a threshold (d_{erreur}) which represents the error-tolerant that the user can accept. Next, for each node in S_1, we compute the distance between it and the centroid. Then we determine the subset S_2 that represents an advanced set of candidate CHs:

$$S_2 = \left\{i \mid i \in S_1,\ d\left(i, \text{centroid}^k\right) \le d_{\text{erreur}}\right\}, \tag{21}$$

where i is a node of S_1 and belongs to the cluster k.

When we select a CH closer to the centroid point, we guarantee that the average distances between the non-CH nodes and the CHs will be minimal:

$$\sum_{i=1}^{|S_k|} d^{\text{LESCA}}_{\text{to CH}_i} \le \sum_{i=1}^{|S_k|} d_{\text{to CH}_i}. \tag{22}$$

Therefore, from (12), we conclude

$$\sum_{i=1}^{N-K} E^{\text{LESCA}}_{\text{nonCH}_i} \le \sum_{i=1}^{N-K} E_{\text{nonCH}_i}. \tag{23}$$

Stage 3. In this stage, each node in S_2 computes its distance to the BS. Then, in each cluster, the nearest node to the BS will

Table 1: Experimental simulation parameters.

Parameter	Value
E_{elec}	50 nJ/bit
ϵ_{fs}	10 pJ/bit/m^2
ϵ_{mp}	0.0013 pJ/bit/m^4
Initial energy E_0	0.5 J
E_{DA} (energy for data aggregation)	5 nJ/bit/message
Area of network	300×300 m^2
BS coordination	(150 m, 350 m)
d_0	88 m
Message size	4000 bits
Number of nodes	500

be elected as a CH. When we select CHs closer to the BS, we guarantee that the average distances between the CHs and the BS will be minimal:

$$\sum_{i=1}^{K} d_{\text{to BS}_i}^{\text{LESCA}} \leq \sum_{i=1}^{K} d_{\text{to BS}_i}. \tag{24}$$

Therefore, from (9), we conclude

$$\sum_{i=1}^{K} E_{\text{CH}_k}^{\text{LESCA}} \leq \sum_{i=1}^{K} E_{\text{CH}_k}. \tag{25}$$

From (23) and (25), we obtain

$$E_{\text{Round}}^{\text{LESCA}} \leq E_{\text{Round}}. \tag{26}$$

We conclude that the total dissipated energy during a round is minimal when the LESCA algorithm is used.

4.4. Data Transmission Step. Based on the id and the numbers of nodes $|S_k|$ in the cluster k, a schedule Time Division Multiple Access (TDMA) will be created automatically to assign to each node a time when it can transmit its data to the CH. If we suppose that the node with id $= i$ is elected as a CH, the node with id $= (i + 1 + |S_k|) \bmod (|S_k|)$ is assigned the first period to transmit its data. Furthermore, we avoid energy consumption due to synchronization of the cluster nodes when the CH is elected to assign the TDMA. To save more energy in a WSN, we assume that if the distance between a node and the BS is lower than the distance between this node and its CH then the node transmits its data directly to the BS.

5. Simulation Results

In this section, we evaluate the performance of the LESCA protocol. MATLAB software was used to simulate its performances. In this study, we have considered first order radio model simulation to LEACH and the simulation parameters for our model are mentioned in Table 1. To validate the performance of LESCA, we simulate a homogeneous clustered WSN in a field with dimensions 300×300 m^2. The total number of sensor nodes $N = 500$. The nodes are randomly

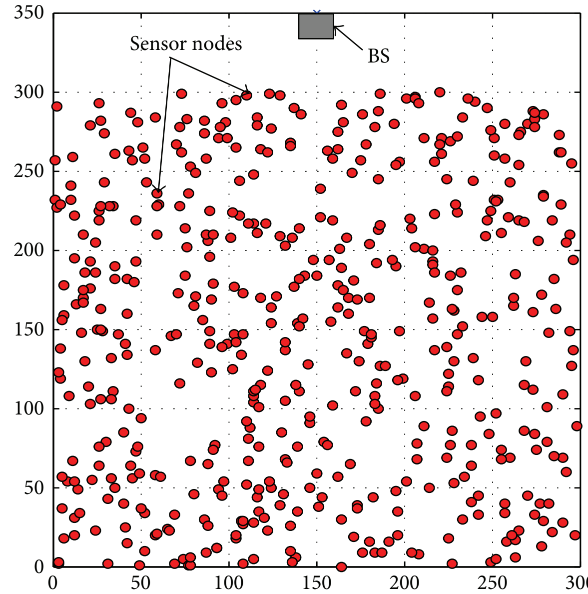

Figure 4: Distribution of wireless sensor network.

distributed over the field. This means that the horizontal and vertical coordinates of each sensor are randomly selected between 0 and maximum value of the dimension. The size of the message that nodes send to their CHs as well as the size of the (aggregate) message that a CH sends to the BS is set to 4000 bits. We compare the performance of our LESCA protocol and other protocols in the same homogeneous setting.

We notice that, for these simulations, the energy of a node decreases each time it sends, receives, or aggregates the data. Besides, each simulation result shown below is the average of 100 independent experiments where each experiment uses a different randomly generated uniform topology of sensor nodes.

In Figure 4, we give an example of WSNs with $N = 500$ nodes randomly distributed in a 300×300 m^2 area.

Figure 5 presents the results of the clustering step using the LESCA algorithm (Algorithm 1). We note that the network is subdivided into nine clusters; this number of clusters is obtained by seeking the maximum difference between two consecutive eigenvalues (Figure 6). As shown, the nodes are correctly distributed over the sensing area. Also, there is no intersection between the different clusters.

In order to evaluate the performances of the new proposed protocol, we propose to compare it to

(i) the LEACH-C algorithm [14],

(ii) the DECSA algorithm [16],

(iii) the spectral classification based on Near Optimal Clustering (Spectral Bipartition) algorithm [20].

The main problem of LEACH-C [14] and DECSA [16] protocols is the random selection of the CH nodes. It is obvious that a stochastic CHS selection will not automatically lead to minimum energy consumption during data transfer for a given set of nodes. All CHs can be located near the

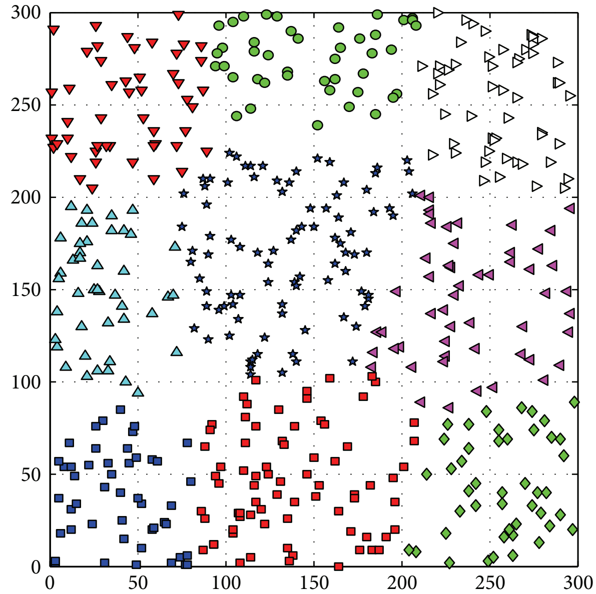

Figure 5: Network clustering using the LESCA algorithm with $N = 500$ and $300 \times 300\,m^2$ area.

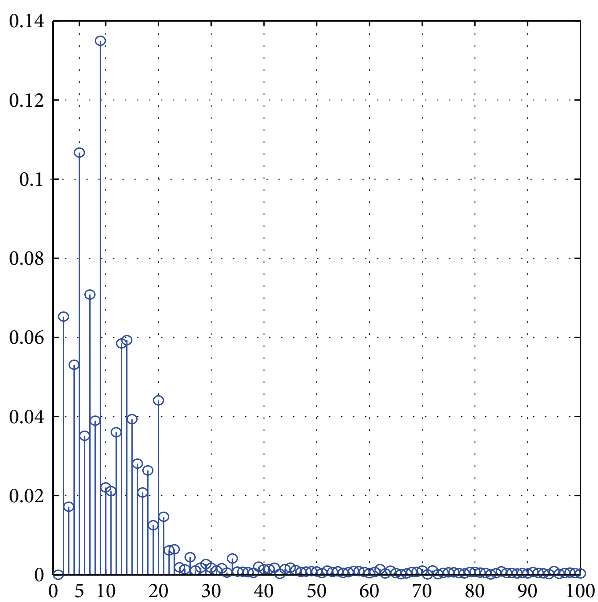

Figure 6: Difference between two consecutive eigenvalues $|\lambda_i - \lambda_{i-1}|$.

edges of the network or adjacent nodes can become CHs simultaneously. In these cases some nodes have to bridge long distances to reach a CH. Also, in Spectral Bipartition algorithm [20], when we use only the second eigenvector of the Laplacian matrix in the clustering setup, we may lose the connectivity information about the network and obtain a poor clustering setup. On the other hand, we cannot partition the network into the optimal number of clusters [22]. The rotation selection of the CHs may obtain a poor cluster head selection setup. Also, when the distribution of CHs is not uniform, some sensor nodes have to transfer data through a longer distance and the reasonable energy saving is not obtained.

We use two metrics to analyze and compare the performances of the protocols.

(i) *Network lifetime*. It can be defined as the time interval between the start of operation (of the WSN) and the death of the first node alive. In this case, we use the first node died (FND) metric. FND denotes the time interval which elapsed until the first node in the network depletes its energy.

(ii) *Number of alive nodes per round*. This measure reflects the total number of nodes that have not yet expended all of their energy. In this case, we use half of the nodes died (HND) metric. HND denotes the time interval which elapsed until half of the nodes in the network are dead.

(iii) *Energy consumption*. Uniform energy consumption is very important for network load balancing. The more uniform energy consumption, the less possibility for node premature death. And the less energy consumption per round, the better network performance.

Figure 7 presents the clustering structure for a certain round in the simulation. According to LESCA features, it is intuitive that each cluster is very well distributed and CHs are located more closely to the cluster centroid.

Figure 8 presents the clustering structure when half of the sensor nodes died. We can show that the loss of half nodes does not automatically diminish the quality of service of the network. Alive sensors are distributed over the coverage area.

Figure 9 gives the curves of the number of nodes alive over time for the Spectral Bipartition, the LEACH-C, the DECSA, and the LESCA algorithms. This figure shows a significant improvement of our protocol in terms of enlarging the network lifetime. Furthermore, it is obvious that the stable time of LESCA is extended for the whole network compared to the other algorithms.

Figure 10 gives the total network energy remaining in every transmission round by using the four compared approaches. This energy decreases rapidly in the Spectral Bipartition, DECSA, and LEACH-C protocols compared to the LESCA one.

Tables 2 and 3 present the performances of the compared protocols using different initial energies of nodes. They give, respectively, the FND and the HND rounds. It is shown that, for different values of the initial energy, our proposed approach presents a significant improvement compared to the others.

Tables 4 and 5 present the effects of the nodes density on the performances of the compared protocols. They give, respectively, the FND and the HND rounds. For different values of N equal to 100, 300, and 500, our algorithm presents an improvement of performances compared to the other algorithms.

Besides, we can justify the robustness of the new approach by the way that the clusters are formed and CH nodes are elected. Indeed, compared to LEACH-C, LESCA considers the residual energy of sensor nodes and the distances between

Given a set of sensor nodes $X = \{x_1, x_2, \ldots, x_N\} \in \mathfrak{R}^{N\times N}$ that we want to cluster into K clusters.
(1) Form the affinity matrix $A \in \mathfrak{R}^{N\times N}$ defined by:

$$A = [a_{ij}] = \begin{cases} \exp\left(\dfrac{-d^2(i,j)}{2\sigma^2}\right) & \textit{if } d(i,j) \le d_0 \textit{ and } i \ne j \\ 0 & \textit{otherwise} \end{cases}$$

(2) Define D to be the diagonal matrix whose (i,i)—*element* is the sum of A's ith row, and construct the matrix $L = D^{-1} \cdot A$
(3) Find $u_1, u_2, \ldots, u_K$ the K largest eigenvector of L (chosen to be orthogonal to each other in the case of repeated eigenvalues) and form the matrix $U = [u_1, u_2, \ldots, u_K] \in \mathfrak{R}^{N\times K}$ by stacking the eigenvectors in columns.
(4) Treating each row of U as a point in $\mathfrak{R}^K$, cluster them into K clusters via K-means.
(5) Finally, assign the sensor node x_i to cluster C_j if and only if row i of the matrix U was assigned to cluster C_j.

Algorithm 1: Clustering step.

Table 2: Impact of the initial energy quantity on the performance of the four compared algorithms (FND) $N = 500$.

Initial energy of nodes	First node died			
	LESCA	Spectral Bipartition	LEACH-C	DECSA
0.5 J	174	120	16	24
1 J	318	253	29	32
1.5 J	445	375	37	51

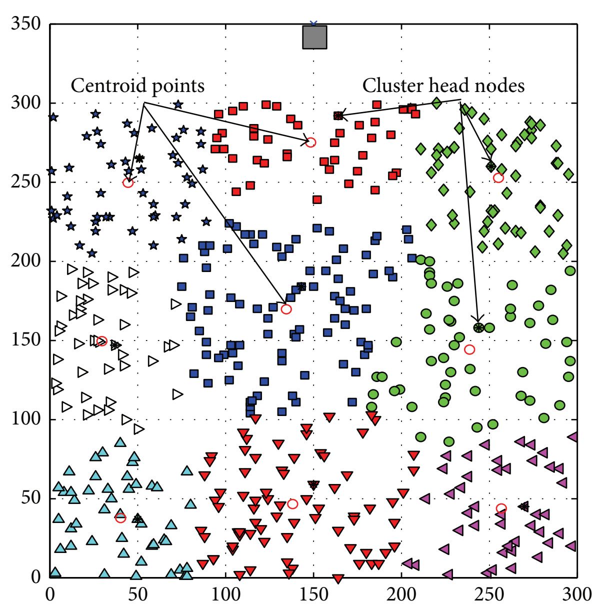

Figure 7: Snapshot of the network when cluster heads are elected.

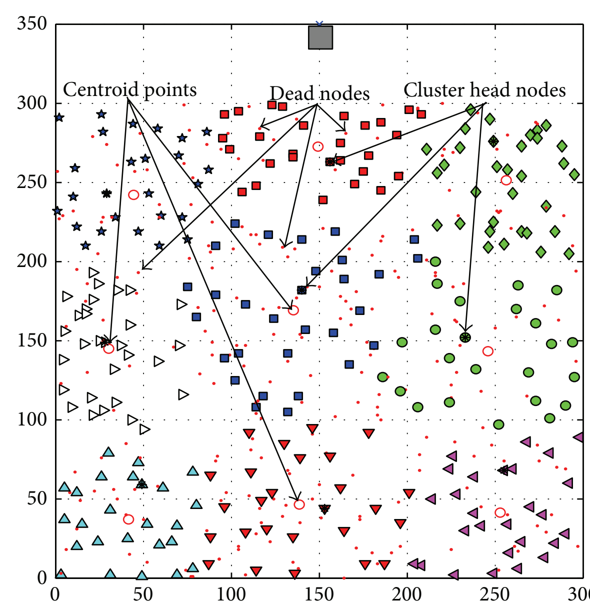

Figure 8: Snapshot of the network when half of the nodes are dead.

sensor nodes and the BS when it chooses the CHs. Also, compared to DECSA, LESCA considers the distances between the centroids points and sensor nodes when it chooses the CHs. These additional two distances allow determining an optimal position of the new CH to be elected. This position guarantees that the consumed energy of the CHs will be more optimal than the energy of the CHs determined by other algorithms. Finally, compared to the Spectral Bipartition, our approach defines an optimal number of clusters by the

Table 3: Impact of the initial energy quantity on the performance of the four compared algorithms (HND) $N = 500$.

Initial energy of nodes	Half of the nodes died			
	LESCA	Spectral Bipartition	LEACH-C	DECSA
0.5 J	559	429	39	130
1 J	1041	857	50	257
1.5 J	1761	1278	75	383

Table 4: Impact of the nodes density N on the performance of the four compared algorithms (FND) $E_0 = 0.5$ J.

Number of nodes	First node died			
	LESCA	Spectral Bipartition	LEACH-C	DECSA
100	77	50	9	18
300	150	115	14	20
500	174	120	16	24

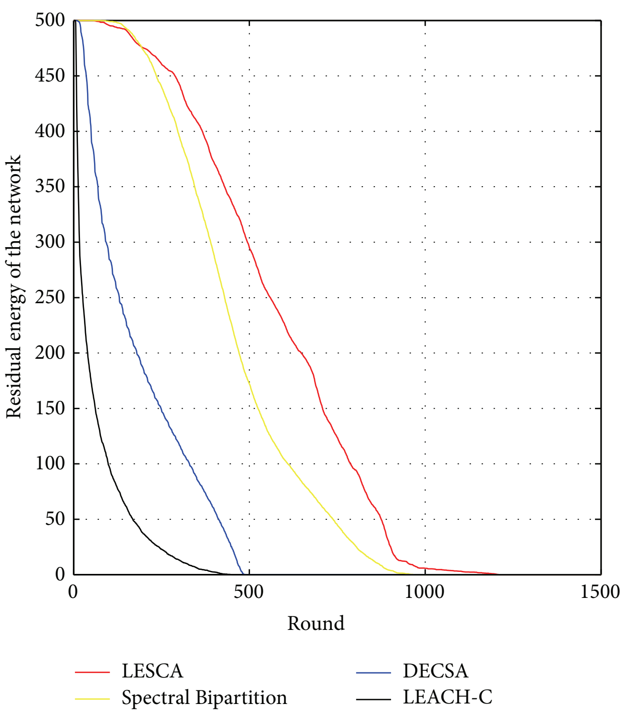

Figure 9: Number of nodes alive over time of the compared protocols $N = 500$.

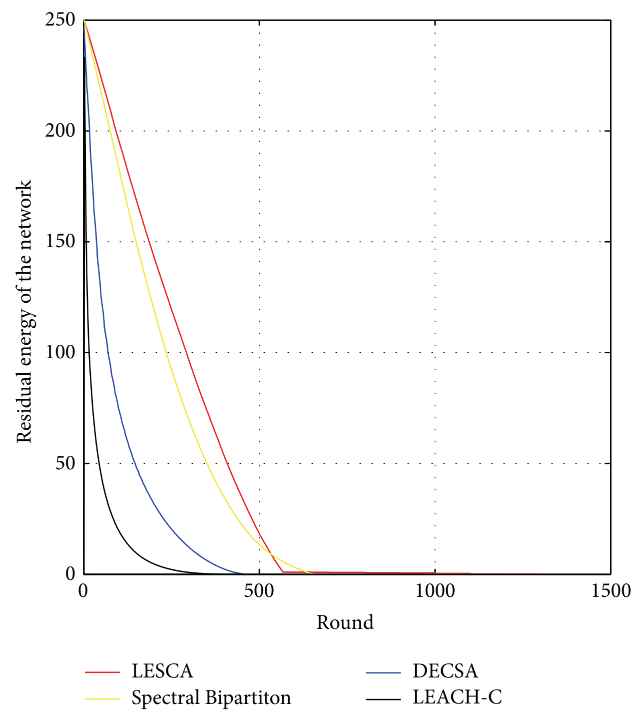

Figure 10: Evolution of the remaining energy in the network during the transmission rounds.

distances between sensor nodes and the spectral classification theory.

6. Conclusion

In this paper we have presented and detailed a new method of clustering in WSN based on spectral classification. Our approach optimizes the energy consumption. In particular, it uses the method of K-ways to cluster the network and a new technique to select CHs. Furthermore, the approach considers both the position of each node and its residual energy to determine an optimal position of the appropriate CHs. Particularly, distances between nodes and the BS and distances between nodes and the centroids of the clusters allow minimizing the consumed energy in the network. In addition, we have measured and compared robustness and performances of our algorithm and three others. The experiments results have shown that LESCA presents a significant performance improvement in terms of energy and lifetime gains, compared to the others. Further works remain for studying other spectral classification techniques which may be more efficient in this kind of applications. Selecting the robust one will be the primordial step in the coming work.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Table 5: Impact of the nodes density N on the performance of the four compared algorithms (HND) $E_0 = 0.5$ J.

Number of nodes	Half of nodes died			
	LESCA	Spectral Bipartition	LEACH-C	DECSA
100	292	124	21	43
300	469	273	26	78
500	174	120	39	130

References

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102–114, 2002.

[2] J. N. Al-Karaki and A. E. Kamal, "Routing techniques in wireless sensor networks: a survey," *IEEE Wireless Communications*, vol. 11, no. 6, pp. 6–28, 2004.

[3] S. H. Gajjar, S. N. Pradhan, and K. S. Dasgupta, "Wireless sensor network: application led research perspective," in *Proceedings of the IEEE Recent Advances in Intelligent Computational Systems (RAICS '11)*, pp. 25–30, Thiruvananthapuram, India, September 2011.

[4] M. F. Othman and K. Shazali, "Wireless sensor network applications: a study in environment monitoring system," in *Proceedings of the International Symposium on Robotics and Intelligent Sensors (IRIS '12)*, vol. 41, pp. 1204–1210, Procedia Engineering, 2012.

[5] D. Estrin, L. Girod, G. Pottie, and M. Srivastava, "Instrumenting the world with wireless sensor networks," in *Proceedings of the IEEE Interntional Conference on Acoustics, Speech, and Signal Processing (ICASSP '01)*, pp. 2033–2036, May 2001.

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

[7] 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.

[8] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 888–905, 2000.

[9] M. E. Newman, D. J. Watts, and S. H. Strogatz, "Random graph models of social networks," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, supplement 1, pp. 2566–2572, 2002.

[10] M. C. Nascimento and A. C. de Carvalho, "Spectral methods for graph clustering—a survey," *European Journal of Operational Research*, vol. 211, no. 2, pp. 221–231, 2011.

[11] U. von Luxburg, "A tutorial on spectral clustering," *Statistics and Computing*, vol. 17, no. 4, pp. 395–416, 2007.

[12] O. Boyinbode, H. Le, A. Mbogho, M. Takizawa, and R. Poliah, "A survey on clustering algorithms for wireless sensor networks," in *Proceedings of the 13th International Conference on Network-Based Information Systems (NBiS '10)*, pp. 358–364, September 2010.

[13] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proceedings of the 33rd Hawaii International Conference on System Sciences (HICSS '00)*, vol. 8, p. 8020, IEEE Computer Society, Maui, Hawaii, USA, January 2000.

[14] 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.

[15] M. Ye, C. Li, G. Chen, and J. Wu, "Eecs: an energy efficient clustering scheme in wireless sensor networks," in *Proceedings of the 24th IEEE International Performance, Computing, and Communications Conference (IPCCC '05)*, pp. 535–540, 2005.

[16] Z. Yong and Q. Pei, "A energy-efficient clustering routing algorithm based on distance and residual energy for wireless sensor networks," *Procedia Engineering*, vol. 29, pp. 1882–1888, 2012.

[17] S. Jerusha, K. Kulothungan, and A. Kannan, "Location aware cluster based routing in wireless sensor networks," *International Journal of Computer & Communication Technology*, vol. 3, no. 5, pp. 0975–7449, 2012.

[18] Z. Rehena, S. Roy, and N. Mukherjee, "Topology partitioning in wireless sensor networks using multiple sinks," in *Proceedings of the 14th International Conference on Computer and Information Technology (ICCIT '11)*, pp. 251–256, Dhaka, Bangladesh, December 2011.

[19] J.-Y. Chang and P.-H. Ju, "An efficient cluster-based power saving scheme for wireless sensor networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2012, article 172, 2012.

[20] B. Elbhiri, S. E. Fkihi, R. Saadane, and D. Aboutajdine, "Clustering in wireless sensor networks based on near optimal bipartitions," in *Proceedings of the 6th Euro NF Conference on Next Generation Internet (NGI '10)*, pp. 1–6, IEEE, Paris, France, June 2010.

[21] J. Shi, "Learning segmentation by random walks," in *Advances in Neural Information Processing*, pp. 470–477, MIT Press, 2000.

[22] A. Y. Ng, M. I. Jordan, and Y. Weiss, "On spectral clustering: analysis and an algorithm," in *Proceedings of the 15th Annual Neural Information Processing Systems Conference (NIPS '01)*, pp. 849–856, MIT Press, December 2001.