Hindawi
Wireless Communications and Mobile Computing
Volume 2017, Article ID 1914956, 14 pages
https://doi.org/10.1155/2017/1914956

Research Article

Evidence-Efficient Multihop Clustering Routing Scheme for Large-Scale Wireless Sensor Networks

Zhihua Li and Ping Xin

Department of Computer Science and Technology, School of Internet of Things Engineering, Jiangnan University, Wuxi, Jiangsu 214122, China

Correspondence should be addressed to Zhihua Li; zhli@jiangnan.edu.cn

Received 22 August 2017; Accepted 30 October 2017; Published 19 November 2017

Academic Editor: Paolo Barsocchi

Energy consumption and transmission reliability are the most common issues in wireless sensor networks (WSNs). By studying the broadcast nature of data transmission in WSNs, the mechanism of guaranteeing reliable transmission is abstracted as propagation of responsibility and availability. The responsibility and availability represent the accumulated evidence of nodes to support reliable transmission. Based on the developed mechanism, an evidence-efficient cluster head rotation strategy and algorithm are presented. Furthermore, backbone construction algorithm is studied to generate the minimum aggregation tree inside the candidate cluster heads. A minimum aggregation tree-based multihop routing scheme is also investigated, which allows the elected cluster heads to choose the optimally main path to forward data locally and dynamically. As a hybridization of the above, an evidence-efficient multihop clustering routing (EEMCR) method is proposed. The EEMCR method is simulated, validated, and compared with some previous algorithms. The experimental results show that EEMCR outperforms them in terms of prolonging network lifetime, improving transmission reliability, postponing emergence of death nodes, enhancing coverage preservation, and degrading energy consumption.

1. Introduction

Decreasing energy consumption, improving energy efficiency, and enhancing transmission reliability are still main challenges of wireless sensor networks (WSNs). The related technique-efficient issues, such as clustering routing, topology control, and multihop transmission, are widely used to improve energy efficiency for WSNs [1–20]. On the other hand, it is very important to note that hybridization combination of various approaches may affect total performance.

Hierarchical topology control, in which nodes are grouped into clusters and cluster heads (CHs) are elected for each cluster to form a backbone construction, can effectively utilize the limited resources of sensor nodes. Hierarchical topology benefits maximizing the network lifetime and optimizing the data delivery ratio on each link. Clustering technique has been proven energy-efficient in WSNs [3, 5–7, 10, 11, 13], in which the low-energy adaptive clustering hierarchy (LEACH) [21] protocol is the most typical one. However, because of its energy-intensive data transmission and routing tasks, a CH node consumes much more energy than regular sensors. Thus, an energy-efficient mechanism for CHs rotation or election, minimizing energy consumption of each node, and maximizing the network lifetime while guaranteeing transmission reliability are still attractive challenges.

Multihop routing is one of two main communication modes for WSNs. In comparison with multipath routing, its transmission has generally been considered an efficient energy-saving approach, especially for large-scale sensor networks [3, 7, 10, 14, 16]. The most commonly used multihop topology is the aggregation tree rooted at the sink [22]. However, the tree topology has an inherent deficiency in that each sensor has only one path to the sink, and the path is not necessarily optimal which results in that the traffic flow passing through these sensors may be unbalanced, thereby making some sensors run out of their energy quickly and even shortening the network lifetime [22]. On the other hand, the clustering routing algorithms [15, 17] integrated the famous Dijkstra algorithm and used the Dijkstra algorithm to build

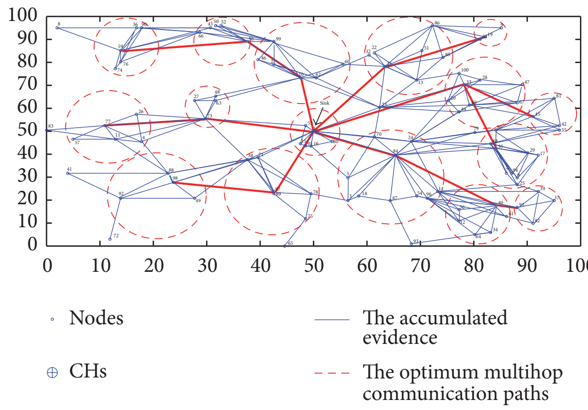


Figure 1: A demo of the principles employed in the proposed scheme.

the shortest path tree with minimum energy consumption for each CH to the sink node. Namely, the CHs could automatically form a number of multihop communication paths. The collected data is continuously transmitted to the sink node via the cluster head-adjacent multihop routing that effectively shared the overloading of different CHs. However, the deficiency of imbalanced distribution of CHs has not been treated, which makes the size of some clusters too large, and it is easy to incur the fast energy consumption in some local clusters.

In this paper, we also address the LEACH-improved scheme. By studying the broadcast nature of data transmission in WSNs, it is abstracted as a propagation of responsibility and availability of nodes. The responsibility and availability integrate sufficient considerations on various network factors including the residual energy of nodes, distance between nodes, distance between the CHs and the base station, and energy loss on the node-joint links. Essentially, the responsibility and availability express the accumulated evidence of nodes to support high-quality communication. Namely, the responsibility and availability represent the comprehensive capability of nodes against node failures and link losses. Based on the above development, an evidence-efficient CH rotation scheme and algorithm are proposed. Furthermore, in order to improve the data transmission reliability between the candidate CHs to the base station, the well-known Kruskal algorithm [23] is employed to generate the minimum aggregation tree; that is, a backbone construction algorithm is also developed. As a hybridization of the presented subalgorithms, an evidence-efficient multihop clustering routing (EEMCR) method is proposed. An example of the working mechanism employed in the EEMCR method is demonstrated by several clusters in Figure 1. And more, the experimental comparison and analysis are also examined with the previous algorithms.

As far as we know, the main contribution of this paper has at least the following points: (1) the eligibility evidence of a node as a CH is regarded as the sum of responsibility and availability (e.g., the accumulated evidence), which increases its robustness against random node failures; (2) the mechanism of clustering is in terms of the propagation of responsibility and availability globally, which reduces the negative intercluster communication interference; (3) the optimum multihop communication paths are achieved by the backbone construction algorithm for the CHs, which ensures the node-joint link reliability against link failures; (4) the associated algorithms and methods are developed and have validated their promising performance.

The rest of this paper is organized as follows. The related works are introduced in Section 2. In Section 3, we describe the system assumptions and communication models. In Section 4, the mechanism and algorithm of CH rotation are examined and proposed. We derive the framework of the EEMCR scheme and present the overview and the detailed design of it in Section 5. Promising experiment results are given in Section 6, and from the effectiveness and efficiency perspective, some validations and comparisons are performed, which are followed by the concluding remarks and future works in Section 7.

2. Related Works

2.1. General Clustering Routing. Grouping nodes into clusters has been the most popular approach for supporting scalability in WSNs [4]. Besides the well-known LEACH and LEACH-developed algorithms, significant attention has been paid to clustering routing mechanism and algorithms yielding a large amount of publications [3, 5–7, 10–13, 17, 18, 20, 24]. These proposed clustering techniques usually include three phases: CH determination, clustering, and data transmission. The previous research works address either one of the issues or overall three phases from different perspectives, and also their improvements and developments with respect to the existing researches.

Usually, clustering is typically based on the energy reserve of sensor nodes and node's proximity to the CH [25]. Regardless of any of their improvements or development, the reduction of energy consumption, prolonging the network lifetime, and improving transmission reliability are the basic goals in WSNs. Energy-efficient clustering technique is applied to reduce energy consumption, interference, and maintaining connectivity and coverage in WSNs. Younis et al. [26] proposed a Hybrid Energy-Efficient Distributed Clustering (HEED) protocol. HEED is a distributed clustering protocol, in which a CH election scheme is presented with the comprehensive treatment of the residual energy and intracluster communication cost. HEED ensures a uniform distribution of CHs and inter-CH connectivity by an adjustable probability of CH election. But the mechanism that the sensor doubles its probability to become CH during a repetition phase is not reasonable enough. Zhou et al. [27] proposed an Energy-Efficient Strong Head clustering (EESH). In EESH, nodes are promoted CHs according to their respective residual energies, their respective degrees, and the distance to and the residual energy of their neighbors. For that, EESH evaluates a cost function for every sensor in the network and iteratively elects the node having the greatest cost as CH. This process terminates when all the sensors in the network are connected to at least one CH. Chamam and Pierre [28] proposed a distributed energy-efficient cluster formation (EECF) protocol. EECF elected the CHs following a three-way message exchange between each sensor and its neighbors. Sensor's eligibility to be elected CH is based on its residual energy and its degree. Thus, the message exchanges complexity of $O(1)$ and a worst-case convergence time complexity of $O(N)$.

Additionally, some nonclustering routing methods or models [29, 30] have also been published. They also have prominent performance on either or all aspects such as degrading excessive communication energy consumption, prolonging the lifetime, and enhancing the transmission reliability.

2.2. LEACH-Improved Clustering Routing. The LEACH protocol, one of the first clustering routing protocols proposed for WSNs, is an adaptive, distributed algorithm that forms clusters of sensors based on the received signal strength and uses local CHs as routers to the sink node [28]. In LEACH, each node has an equivalent opportunity to become CH, and through a random rotation of CHs, LEACH provides a balance of energy consumption for each node. However, CHs transmit data directly to the sink node, which can be energy-consuming in large-scale WSNs. Power-efficient Gathering in Sensor Information Systems (PEGASIS) [31] and Hierarchical-PEGASIS are two improvements of LEACH. Unlike the existing multiple clusters in LEACH, PEGASIS and Hierarchical-PEGASIS constructed chains of sensor nodes so that each sensor node is transmitted and received from a neighbor and only one node was selected from that chain to transmit data to the sink node [28]. Unfortunately, the communication between the elected CH and the base station is one-hop, which may waste energy and prove to be unsuitable for large-scale WSNs [28]. Obviously, they need to be improved in-depth.

Based on the LEACH algorithm, two factors, the energy and distance, were cast into modifying the threshold function of LEACH protocol; a cluster head multihop routing improved algorithm (CMRAOL) based on LEACH [21] was proposed, in which the multihop communication approach was adopted and a reliable path was built between the CHs and the sink node; the energy consumption of the network was effectively balanced, but the CH election process was too complex. Thus, the algorithm was only suitable for static networks; the improved low-energy adaptive clustering hierarchy-centralized (LEACH-C) algorithm [13] sufficiently considered two factors: the energy and number of CHs. The energy consumption of each node improves to be more balanced, but the clustering overhead was too large.

Additionally, some of the existing clustering routing algorithms [3, 7, 10, 13, 14] did not consider the residual energy and the deployment location of nodes; in response to this deficiency several researches are carried out in the EBAPC algorithm [18]. The EBAPC algorithm defined a new concept of fitness factor and employed the strategy of cluster center determination in AP algorithm [24]. The cluster center determination strategy [24] was approximated as an election scheme of the CHs with sufficient consideration of the residual energy of nodes. In contrast to the previous algorithms, the CH election was more reasonable and the energy consumption got more balanced. However, the EBAPC algorithm did not regard the relation between the location of the base station and the entire energy consumption of the network, which easily leads to high energy consumption of some local nodes; even premature death of some important nodes affects the network lifetime and the service quality of the network; the ELBC and BM-ELBC algorithm [19] based on EBAPC algorithm were proposed, respectively. A new concept of energy level with sufficient consideration of two factors, the residual energy of nodes and distance between different nodes, was introduced; the proposed algorithms made energy consumption more balanced, and they significantly prolonged the network lifetime. However, the definition of energy level was relatively rough, which easily spurred an unreasonable CH election such as taking the remotely unreasonable nodes as the second level CHs. This situation ultimately affected the transmission efficiency and transmission delay, as well as the network lifetime. A clustering routing algorithm that introduces a new definition of node competitiveness based on it was presented [20]. The APBCS algorithm made the node-joint clustering more uniform and made the node deployment more reasonable. However, the APBCS algorithm had not given an appropriate solution scheme to satisfy the requirement of multihop transmission scheme within the node-joint CHs, as a result, increasing the energy consumption on some CHs. When the scale of WSNs gets large, it leads to the occurrence of a premature death for the CHs, such that it was unable to guarantee the coverage preservation and network connectivity.

The differences covered between the proposed EEMCR scheme and the previous researches include the following: the nature of data transmission in WSNs is abstracted as propagation of responsibility and availability, and the responsibility and availability are taken as the accumulated evidence of nodes for guaranteeing transmission reliability; the responsibility and availability integrate comprehensive considerations on several network factors; an evidence-efficient cluster head rotation mechanism and algorithm are presented; backbone construction algorithm is developed to generate a minimum aggregation tree; essentially, the branches inside the generated minimum aggregation tree construct the multihop communication paths. The experimental results show that there exists a remarkable improvement on prolonging the network lifetime, postponing the death of nodes, saving energy, and conducting coverage preservation.

3. Assumptions and Models

3.1. System Assumptions. Assuming that n nodes are randomly deployed in a target area, in which each node has the same configuration and data processing capability, the location of each node is known-available. Additionally, this paper makes the following assumptions:

(1) All the sensor nodes in the system have the same organization and initial energy E_0.

(2) All sensor nodes adaptively adjust the transmission power according to their need.

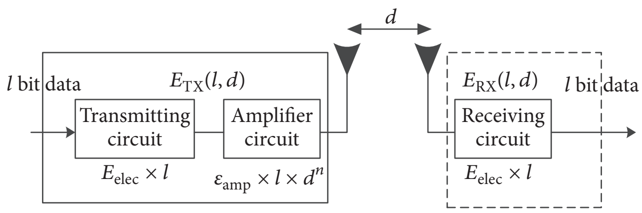


Figure 2: Wireless transceiver circuit energy consumption model.

3.2. Communication Model. In this paper, a well-known common first-order wireless energy consumption model [32] is employed, which is shown in Figure 2.

With respect to the above model, the energy consumption of transmitting l bit data is composed of transmission circuit power and power amplification losses. Different power consumption models are employed with different distances; and the energy consumption of transmitting l bit data by nodes can be calculated by the following formula:

$$E_{TX}(l,d) = ETX_{\_elec}(l) + ETX_{\_amp}(l,d) = \begin{cases} lE_{elec} + l\varepsilon_{fs}d^2 & d < d_0 \\ lE_{elec} + l\varepsilon_{amp}d^4 & d \geq d_0, \end{cases} \quad (1)$$

where E_{elec} represents the energy consumption by a circuit processing a single bit data, ε_{fs} denotes the coefficient of power amplifier in the free space model, ε_{amp} is the coefficient of multidiameter attenuation model of power amplifier, and d_0 is the critical distance between free space propagation model and multipath attenuation model and is calculated as the following formula:

$$d_0 = \sqrt{\frac{\varepsilon_{fs}}{\varepsilon_{amp}}}. \quad (2)$$

The energy consumption by the node receiving l bit data is formalized as the following formula:

$$E_{Rx}(l) = E_{Rx-elec}(l) = lE_{elec}. \quad (3)$$

4. Evidence-Efficient Cluster Head Rotation Mechanism

In this section, we discuss the issue of CH rotation. First, we think that the essence of CH election in clustering topology is equivalent to determining the cluster centroids of cluster algorithms in data mining. The Affinity Propagation (AP) [24] cluster algorithm is very prominent at its performance on carrying out large amount of data. Thus we employ the propagation mechanism in AP algorithm and study the scheme of CH rotation associated with it. Under sufficient consideration of several network factors including the residual energy of nodes, distance between nodes, energy consumption associated transmission distance and amount of data, and abstracting the capability of guaranteeing transmission reliability in WSNs as the responsibility and availability propagation between node-joint links, the maximum sum of responsibility and availability is taken as the accumulated evidence of CH rotation. Namely, if the accumulated evidence of a node is relatively large, then it has a relatively high possibility of becoming a CH, which is in charge of maintaining the node-joint links and ensuring transmission reliability.

Additionally, since the energy efficiency is one of the critical factors that influence the network lifetime, in particular, several new definitions with respect to the measurement or estimation of energy consumption are presented as follows.

Definition 1. Node-Energy-Level (NEL) is used to measure the energy consumption of nodes in the network. NEL is calculated as Formula (4).

$$\text{NEL} = \left(\frac{E^{\text{remain}}}{E^{\text{init}}}\right)^{1/N_{\text{live}}}, \quad (4)$$

where N_{live} denotes the number of surviving nodes in the current status of the network, E^{init} represents the initial energy of a node, E^{remain} denotes the residual energy of a node. The meaning of NEL shows that the higher the residual energy of a node, the greater the value of the node energy level (e.g., NEL), which indicates that the stronger the current node activity, the greater the coverage preservation as well as the better the quality of service (QoS). In contrast, when the residual energy of nodes is lower, the node energy level is smaller, which indicates that the smaller coverage preservation and the worse QoS.

Definition 2. The Energy-Cost EC(i, j) is defined as the following formula:

$$\text{EC}(i,j) = \frac{\text{ET}(i,j)}{E_i^{\text{remain}}}, \quad (5)$$

where E_i^{remain} denotes the current remainder energy of CH i and ET(i, j) represents the real amount of energy consumed by CH i to transmit a unit data to CH j. EC(i, j) only expresses a logical transmission overhead, rather than a real metric of energy consumption. Its meaning not only shows considering the energy consumption of the real data transmission, but also takes into account the residual energy level of nodes themselves.

Definition 3. $s(i, j)$ indicates the level of node j appropriate for the cluster head to node i, which is called the orientation-tended matrix and defined as the following formula:

$$s(i,j) = \begin{cases} -\left(\alpha * \left(d(j,B) * E_j^{\text{remain}}\right)^{\text{NEL}} + (1-\alpha) * d(i,j)\right), & i \neq j, \alpha \in [0,1] \\ P(j), & i = j, \end{cases} \quad (6)$$

where B represents the base station; $d(i, j)$ is the distance from node i to node j and with the Euclidean distance to express; $d(j, B)$ is the distance between node j and base station; E_j^{remain} is the current residual energy of node j; NEL is the Node-Energy-Level (e.g., NEL) of the current node and is calculated using Formula (4); α is a weight used to adjust the residual energy and distance in the CH election scheme. Usually, a node, with a larger value of α and more remainder energy and closer to the base station, is easier to become a CH. P is the orientation-tended parameter, that is, the diagonal value in the orientation-tended matrix, defined as the following formula:

$$P(i) = -\text{NEL} \times \left(\frac{E_i^{\text{init}}}{E_i^{\text{remain}}} \right), \tag{7}$$

where E_i^{init} is the initial energy of node i. It can be seen from Formula (7) because the orientation-tended matrix is always negative. Thus, the more the residual energy of the node, the greater the value of the orientation-tended parameter P, as well as the higher the probability that the node will become the CH.

Based on the above, this paper integrates inspirations on the broadcast nature of data transmission in WSNs and the idea in AP algorithm. Furthermore, we give improvement recognition of the responsibility and availability, respectively.

Definition 4. The eligibility evidence of a node as a cluster head is quantified using the responsibility and availability. They are defined as follows:

$$r(i,k) = s(i,k) - \max\left\{a\left(i,k'\right) + s\left(i,k'\right)\right\} \quad \left(k' \in \left\{1,2,\ldots,n,k' \neq k\right\}\right), \tag{8}$$

$$a(i,k) = \min\left\{0, r(k,k) + \sum_{i'} \max\left\{0, r\left(i',k\right)\right\}\right\} \quad \left(i' \in \left\{1,2,\ldots,n; i' \neq i; i' \neq k\right\}\right), \tag{9}$$

where $r(i,k)$ is the responsibility between node i and node k. If k is taken as the potential CH, then responsibility is transmitted from node i to node k and is equivalent to the accumulated evidence for node k to be the CH of node i. $a(i,k)$ as the availability between node i and k transmits from node k to node i. It implies the accumulated evidence for node i to select node k as its CH. Thus, it measures whether k can finally become the real CH after each cycle of self-adaptation.

Formulas (8) and (9) iterate according to Formulas (10) and (11), respectively.

$$r^{\text{new}}(i,k) = (1-\lambda)\, r(i,k) + \lambda r^{\text{old}}(i,k), \tag{10}$$

$$a^{\text{new}}(i,k) = (1-\lambda)\, a(i,k) + \lambda a^{\text{old}}(i,k), \tag{11}$$

where $\lambda \in [0,1]$ is the learning rate in updating $r(i,k)$ and $a(i,k)$. At the beginning of the iteration, their initial values are set to "0".

Usually, link losses and node failures are the primary reasons to influence transmission reliability in WSNs; unfortunately, they are negatively affected by many network factors. In this paper, we take the availability and responsibility as the comprehensive evidence against node failures and link losses, which aims to improve the total performance including transmission reliability and other properties. To this point, the availability and responsibility provide a logical metric for nodes to guarantee transmission reliability.

Furthermore, in order to ensure the rationality of CH rotation and correctness of CH election mechanism, we give the following theorem and proof.

Theorem 5. *The greater the sum of responsibility and availability, the greater the possibility of a node to become a cluster head.*

Proof. The essence of a "cluster" in the clustering routing is similar to the "cluster analyze" in the associated cluster algorithms of data mining, thus the CH election likes the procedure of determining the cluster centroids in the algorithms of data mining. Moreover, the well-known AP algorithm is an essential cluster algorithm, and its cluster procedure is carried out by the propagation of responsibility and availability, in which the first critical business is to adaptively determine the different cluster centroids according to the changeable orientation parameter. The cluster procedure is iteratively ongoing until all of the data items are clustered. This above idea is equivalent to the mechanism of the CH rotation in WSNs and equivalently applied in selecting the CHs; that is, there exists a set of dynamic candidate CHs composed of a large amount of nodes, which correspond to the cluster centroids and are traversed by the propagation of responsibility and availability. In each cycle of the clustering, the node-joint links build the interpath within a single cluster and the cluster head-joint links set up the intrapath between different clusters. During the propagation processing, the availability and responsibility denote the capability of resisting nodes failures and link losses, as well as their propagation direction which guides the data transmission paths. Obviously, the greater the sum of responsibility and availability is, the greater the probability that a node becomes a CH would be. The theorem is proved. □

Consequently, in a real clustering phase, each cycle includes two phases: cluster head election and clustering. The two phases are periodically ongoing until any node runs out of its energy. The responsibility and availability transmit themselves within nodes-joint links in the clustering procedure. The sum of "$r(i,k)+a(i,k)$" is adaptively changing. The larger the sum of $r(i,k)+a(i,k)$, the greater the probability that the node k becomes a CH; otherwise, the node k is unable to become a CH. Namely, as for node i, it always selects node k that maximizes the sum of $r(i,k)+a(i,k)$. Furthermore, if $i = k$, node i is the CH; otherwise, $i \neq k$, node k is the CH of node i. In combination with Definitions 1, 2, and 3 and the above-mentioned analysis, it indicates that the nodes with more residual energy and little amount of data to transmit have a high probability to be CHs. Apparently, the CHs election is an adaptive and dynamic process.

Input: Max iterations: *IterMax;*
Output: Candidate cluster head nodes;
Calculate NEL according to *Formula. (4)* and send the NEL value to sink node;
Sink node calculate the S and P according to *Formula. (6) & (7)* respectively;
Sink node broadcast the S and P;
Set $Max = 0$;
while (Current iterations $Iter < IterMax$)
    Calculate $r(i, j)$ and $a(i, j)$ according to *Formula. (8) & (9) & (10) & (11)*;
        Obtain node k with $\max(a(i,k) + r(i,k))$ for node i;
        if $i = k$
            Node i is the cluster head;
        else if $i \neq k$
            Node k is the cluster head of node i;
        end
        Current iterations $Iter = Iter + 1$;
end

Algorithm 1: Clustering.

Lemma 6. *The propagation direction of responsibility and availability, as the multihop transmission path, effectively avoid the node failures.*

Proof. Use reduction to absurdity to prove it. From the perspective of the definition of responsibility and availability, as long as a node fails in the network, its availability automatically gets "0" and the associated responsibility of it also becomes "0"; that is, the responsibility or availability does not continuously spread along such a path including these nodes. Thus the lemma is proved. □

Lemma 7. *The termination condition for the propagation of responsibility and availability is the energy exhaustion of any node in the network.*

Proof. As it can be seen from the aforementioned Definitions 1, 2, 3, and 4, once the energy of a node in the network gets "0", then its corresponding orientation-tended parameter becomes "0". As long as the orientation-tended parameter is "0", the node is unlikely to be elected as the CH in the next cycle of clustering. Similarly, we can see that once the energy of all nodes in the network gets "0", the transmission of the responsibility and availability is automatically terminated.

Evidently, the aforementioned description is a global election scheme for CH rotation. Namely, all of the running nodes also have the same opportunity to become candidate CHs regardless whether they used to be CHs or not. Undoubtedly, this scheme is helpful to achieve balance in energy consumption of each node and prevent the premature death of any node, thereby guaranteeing coverage preservation and prolonging the network lifetime. □

5. EEMCR Routing Scheme

In this section, we talk about the proposed EEMCR routing scheme in detail. Exactly, EEMCR scheme inherits the basic framework of LEACH algorithm, which hybrids the basic processing of clustering and data transmission within each cycle. In fact, EEMCR scheme integrates several subalgorithms on CH rotation and backbone construction; such hybridization subalgorithms are ongoing in an iteration mode and they finally complete multihop clustering routing. The details of each subalgorithm are discussed in the following parts, respectively.

5.1. Clustering.

5.1. Clustering. The EEMCR conducts the clustering in the mechanism of iteration cycle by cycle, which is similar to that in LEACH algorithm. However, the LEACH protocol randomly replaces the CHs in each cycle of iteration; in contrast, the EEMCR method adaptively updates the CHs according to the accumulated evidence of each node. In the initial stage of each cycle for the proposed EEMCR, the base station calculates the orientation-tended matrix of the nodes in terms of the Formula (6) and broadcasts it to all nodes. The nodes that successfully receive the orientation-tended matrix calculate their own responsibility and availability according to Formulas (8) and (9). At the same time, for any node i, the algorithm takes the node k that maximizes the sum of $r(i,k) + a(i,k)$ as the CH in the current cycle of iteration. Along with the ongoing iteration, each node updates its responsibility and availability according to Formulas (10) and (11) until the demands of iterations or the convergence of the algorithm is achieved.

Summarizing the above steps describes the clustering algorithm, as shown in Algorithm 1.

Theorem 8. *The time complexity of clustering algorithm is* $O(n)$.

Proof. The termination condition of Algorithm 1 is either the number of iterations exceeds the set threshold or the CHs without any ongoing changes. Obviously, there is no negative impact on the complexity of the algorithm. Therefore, the time complexity of Algorithm 1 is also $O(n)$. □

```
Q← ∅ //∗ Initialization, Q is the set of the minimum aggregation tree
for each path e(i, j) ∈ G //∗G represents the node set of the candidate cluster heads
    Sort the path e(i, j) into an increasing order by weight EC(i, j);
end
for each e(i, j) ∈ G, taken in an ascending order by weight EC(i, j);
    if e(i, j) ∉ Q&&No loop then
        Q← Q ∪ e(i, j);
    end
end
return Q
```

Algorithm 2: Backbone construction.

5.2. Backbone Construction. Backbone construction essentially consists of generating minimum aggregation tree and minimum aggregation tree-based multihop communication paths according to graph theory. On the other hand, the multihop transmission path selection problem is very complex. The following theorem is given firstly.

Theorem 9. *Multihop routing with respect to the cluster heads in WSNs is an NP-complete problem.*

Proof. At the stage of building the minimum aggregation tree, the energy consumption of each CH depends on its EC (e.g., Definition 2) with its neighbor CHs. It can be seen from Definition 2 that the EC of a CH is determined by its residual energy and its distance to the neighbor-adjacent CHs. Therefore, in order to minimize the energy consumption and improve the transmission reliability, we need to find a spanning tree so that the number of its neighbor CHs for each CH is the smallest and the interdistance between the different CHs is the shortest. Namely, it is necessary to find a minimum spanning tree. Garey and Johnson [33] have proved that the minimum spanning tree problem is an NP-complete problem.

Additionally, the well-known Kruskal algorithm in the field of graph theory is suitable for solving the NP-complete problems [23]. In this paper, we also employ it to conduct the issue of multihop routing selection within CHs to try to achieve an approximate solution for it. In fact, the goal is to construct the connected graph with Kruskal algorithm, in which, the selected child nodes from the candidate CHs take the value of the EC as the vertexes (e.g., Definition 2) between a pairwise of the selected nodes and the weight on the corresponding edge. Finally, the essence of the multihop communication path construction is to generate an aggregation tree from the connected graph, of which the edge with a relative minimum weight is the criteria for selection, namely, selecting the edges from the graph to build a tree according to the ascending order of the weight. Once a single edge with a relatively small weight is selected and the currently selected edge with the previously selected edges does not form a loop, then it is preserved in the generated tree; otherwise, the latest selected edge is removed. As a result, it finally obtains a tree with $n_{\text{cluster}}-1$ items of edges. Thus the theorem is proved. □

Furthermore, the construction of multihop communication paths within the candidate CHs is as follows: first, using the proposed Algorithm 1 to determine the candidate CHs; second, the selected CH nodes deliver a status message to the base station. The status message is expressed as a frame of {*node ID*, *node residual energy*, *location*}; third, the base station calculates the EC between the CHs based on the received status messages and broadcasts the obtained EC information to all of the CHs; finally, the CHs that have received the broadcasting message from the base station cooperatively construct an aggregation tree whose root node is the base station. Consequently, the data can be transmitted to the base station with a multihop transmission approach along the paths inside the generated aggregation tree; the forwarded data get to the base station along the branches from the leaves to the root inside the generated aggregation tree. Essentially, the built aggregation tree is the backbone construction.

In summary, backbone construction algorithm is presented, as shown in Algorithm 2.

Theorem 10. *The generated tree, conducted by the Kruskal-based method, is a minimum aggregation tree.*

Since the typical Kruskal algorithm targets to find an undistorted minimum aggregation tree, thus we use the reduction to absurdity to prove the Theorem 10.

Proof. Assuming that there exists a really minimum aggregation tree T_1 in the network, in contrast, T_2 is another aggregation tree obtained by the Kruskal algorithm and $T_1 \neq T_2$; thus there is at least one path e that falls inside T_2 rather than T_1. If the path e is removed from T_2, T_2 becomes two parts of nonconnected subtrees which are denoted by T_a and T_b, respectively. On the other hand, there must be a path f existing in T_2, of which one of the vertex of f is located in T_a and the other one is located in T_b, satisfying the condition $\text{EC}(f) < \text{EC}(e)$. Furthermore, it is impossible to select path e due to the $\text{EC}(f) < \text{EC}(e)$; in contrast, the path f is preferably selected according to the Kruskal-associated aggregation tree generation method. Under the worst situation, if the path e is forced to join the connected subsets, the loop inevitably emerges. This result is in contradiction with the basic idea of

```
Initializing network parameters;
Sink node collects all the related parameters of the nodes in the network;
if sink node receives all the relevant parameters of the nodes in the network;
    call Algorithm 1
    call Algoithm 2
end
Sink node broadcasts cluster head information and minimum aggregation tree information
for each node i received the cluster head information and minimum aggregation tree information
    if (node i is the CH)
        Find the next hop (e.g., cluster head) from the clues in the minimum aggregation tree;
    else
        Find its cluster head from the candidate cluster heads;
    end
end
for each member of a cluster in the network
    Performing data transmission by single-hop mode;
end
for each CH in the network
    Performing data forwarding;
end
```

Algorithm 3: EEMCR method.

the Kruskal algorithm. Namely, only when $T_1 = T_2$ holds, Theorem 10 is correct. Proof is completed. □

Based on the above analysis and proof, under the consideration of building a new tree denoted as T_3, T_3 satisfies the following formula:

$$T_3 = T_2 - e + f \tag{12}$$

which implies that using path f connects T_a and T_b. In contrast, T_3 is the finally generated aggregation tree using the Kruskal-based method. It is easy to find that the EC of T_3 is less than that of T_2, which implies retaining the edge with a small weight of EC and removing the edge with a large weight of EC. Similarly, as for m different paths which exist in T_1 and T_2, the minimum aggregation tree T_1 can be successfully achieved by m times of the transformations, like the above steps. Namely, the agglomeration trees in the network can be transformed into the minimized aggregation tree by the Kruskal-like algorithm; that is, the minimum aggregation tree T_1 must be available.

5.3. Evidence-Efficient Multihop Clustering Routing Scheme. Hybridizing the aforementioned Algorithm 1, Algorithm 2 and some necessary steps are described as an evidence-efficient multihop clustering routing (EEMCR) method, as shown in Algorithm 3. In Algorithm 3, to the intracluster communication, the CH and its members communicate directly with single-hop transmission mode, whereas the multihop communication approach is examined within interclusters.

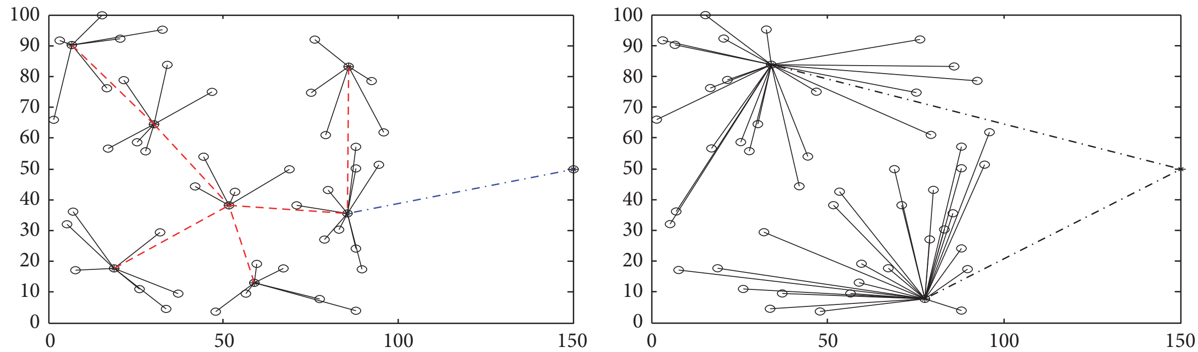
Theorem 11. *The time complexity of EEMCR scheme is $O(n\log n)$.*

Proof. Assuming the connected graph $G(V, E)$ with n_t vertices and n_e edges, using results from the process of building the minimum aggregation tree according to the Kruskal algorithm, and the weight of each edge is the corresponding EC (e.g., Definition 2). Firstly, since the Kruskal algorithm builds a subgraph with only n_t vertices without any edge, at this time, this subgraph can be equivalently regarded as a forest with n_t items of trees, and the vertices in the subgraph are taken as the root node of each tree, respectively. Then, the Kruskal algorithm selects an edge from the candidate set of edges in the network with the relative smallest weight, and if the two vertices of the selected edges belong to two items of different trees, the two pieces of trees are combined as a new tree by connecting the two vertices; that is, the selected edge is added to the subgraph. In contrast, if the two vertices of the selected edge have fallen into the same tree, the selected edge is not desirable; instead, the next edge with the relative smallest weight is tried continuously. This similar process continues until there remains only one tree in the forest; that is, the built subgraph contains $n_t - 1$ items of edges. Consequently, EEMCR scheme just scans the n_e items of edges at most once, and selecting the edge with the minimum EC requires only the time of $O(n\log n)$. Thus, the time complexity of Kruskal-based minimum aggregation tree generation method is $O(n\log n)$.

In summary, since the time complexity of clustering is $O(n)$ and the time complexity of backbone construction is $O(n\log n)$, the time complexity of EEMCR method is $O(n\log n)$. □

6. Experimental Analysis and Comparison

In this section, several experiments are arranged to validate the effectiveness and efficiency of the proposed method from different aspects. The parameters configured in the simulation experiments are shown in Table 1. All of the experiments are realized using Matlab R2010b.



(a) The clustering results using EEMCR

(b) The clustering results using LEACH

Figure 3: The clustering results.

Table 1: Parameters in the experiments.

Parameter type	Parameter value
Node distribution area	100 m × 100 m
Sink location	(150, 50)
Node numbers	100~850
Initial energy of sensor node E_0	0.3 J
Circuit processing data consumption energy E_{elec}	50 nJ/bit
Packet length	2000 bits
Control packet length	32 bits
ε_{fs}	100 pJ/bit/m^2
ε_{amp}	0.0013 pJ/bit/m^4
Data aggregation energy consumption	5 nJ/bit/signal
Data aggregation rate	0.6
λ	0
IterMax	1000
γ	0.9

6.1. Effectiveness. Figures 3(a) and 3(b) are the experimental results of clusters using EEMCR and LEACH algorithms, respectively. The horizontal and vertical coordinates of the graph represent the location information of the nodes, “0” represents an ordinary node, and the sink’s location is (150, 50). The noncluster head nodes in Figure 3(a) are connected to the CH with a red dashed line, and the CH node is connected to the sink with a dark blue dashed line; the noncluster head node in Figure 3(b) is connected to the CH with a solid line, and the CH is connected to the sink with a dashed line. Obviously, due to EEMCR scheme taking into account several network factors, the distribution of the determined CHs is reasonable and the scale of clusters is relatively appropriate; in contrast, the CH distribution using LEACH algorithm is random and the cluster size is too large.

Furthermore, to validate the influence of the proposed clustering and backbone construction algorithm on the performance, we fundamentally implement the LEACH [21], LEACH-C [13] and CMRAOL [14] algorithms, respectively. Additionally, through casting the proposed CH rotation mechanism and minimum aggregation tree generation instead of the original ones in the LEACH algorithm, the hybrid results are expressed as LEACH-CS and LEACH-MT, respectively. Namely, embedding the proposed clustering algorithm replaces the clustering phase in LEACH and generates the LEACH-CS method; the backbone construction algorithm builds the communication path instead of the original one in LEACH; the generated combination method is denoted as LEACH-MT. So, the LEACH, LEACH-C, and LEACH-CS are associated with the LEACH algorithm and with or without considering the proposed cluster head rotation mechanism, whereas the LEACH-MT and CMRAOL algorithms consider the multihop transmission approach. Finally, the effectiveness is validated on several indices including the network lifetime (LifeTime), emergence of first death node (FirstDeath), maximum average energy consumption (MaxEC) and minimum average energy consumption (MinEC) of nodes. The experimental results are shown in Table 2.

Table 2: Comparisons of the five algorithms.

Name	FirstDeath	LifeTime	MaxEC	MinEC
LEACH	24	301	0.0167	0.0011
LEACH-CS	65	377	0.0074	0.0015
LEACH-C	43	345	0.0141	0.0007
LEACH-MT	36	326	0.0117	0.0009
CMRAOL	33	313	0.0150	0.0021

Regarding the first emergence of death nodes in Table 2, the first emergence time is during the 65th cycle of clustering in LEACH-CS that is far later than that of the other compared algorithms; LEACH-C is the second best, wherein the LEACH algorithm first causes the death node to emerge at its 24th cycle and is the worst. This is because LEACH-CS employs the proposed evidence-efficient cluster head rotation mechanism to elect the CHs and achieves a significantly delayed energy consumption of each node and effectively prolongs the network lifetime. Although the LEACH-C algorithm takes the energy factor into account in the CH election stage, it consumes more energy in the clustering stage, which leads to the premature death of the cluster nodes compared to the LEACH-CS. However, the first emergence of death nodes in LEACH-C compared with that in LEACH algorithm is improved; the cause reason is because LEACH-C utilizes a cluster head election strategy against that of randomly selecting the CHs in LEACH algorithm. The improper mechanism in LEACH results in insignificantly large amounts of energy consumption.

The network lifetime mainly reflects the capability of different algorithms to allocate the energy and schedule the task of data transmission. As shown in Table 2, the network lifetime of LEACH-CS algorithm is the longest and the network lifetime of LEACH-C algorithm is the second longest. The network lifetime of these two algorithms greatly improved in comparison with that of LEACH algorithm. The main reason is that a relatively proper cluster head election is taken in them, which results in slowing energy consumption and effectively prolonging their lifetime. On the other hand, the lifetime of LEACH-MT and CMRAOL algorithms is relatively close and also longer than that of the LEACH algorithm. The primary reasons are that the LEACH-MT and CMRAOL employ a multihop communication mode and leverage link redundancy to improve data transmission efficiency and energy consumption. However, in comparison with the LEACH-CS and LEACH-C, the LEACH-MT and CMRAOL only lead to a little achieved improvement for the network lifetime since the CH election is not entirely reasonable in the clustering stage, in which the clustering process still consumes a large amount of energy against the LEACH-CS algorithm.

The maximum average energy consumption and minimum average energy consumption of the nodes is analyzed. The difference between the maximum average energy consumption and minimum average energy consumption for the nodes is regarded as the index, a larger difference indicates the more energy consumption is centralized on a few nodes with a large amount of workload; in contrast, a smaller difference indicates the energy consumption and its distribution is more reasonable and uniform, which benefits the coverage preservation. From Table 2, the difference of average energy consumption using LEACH-CS is the smallest one of 0.0059; the difference of average energy consumption using LEACH algorithm is the largest one of 0.0156. Although the LEACH-MT and CMRAOL algorithms employ multihop communication approach, their difference of average energy consumption also reaches more than 0.01. The above results indicate that LEACH-CS achieves a promising balance between energy consumption of nodes, thus prolonging the network lifetime and enhancing coverage preservation.

In summary, the hybridization LEACH-CS method shows the best performance in terms of the indices LifeTime, FirstDeath, MaxEC, and MinEC. The experimental results indicate that the proposed cluster head rotation mechanism effectively limits the transmission range and set of neighbors of nodes.

6.2. Efficiency. In this part, experimental comparisons and analyses are performed between the EEMCR method and other methods, such as the LEACH-developed algorithm including BM-ELBC [19], ELBC [19], and EBAPC [18], as well as LEACH [21] itself, on the indices including the total energy consumption of networks, survival time of nodes, average energy consumption of nodes, and influence of the sink location, which are to show its efficiency.

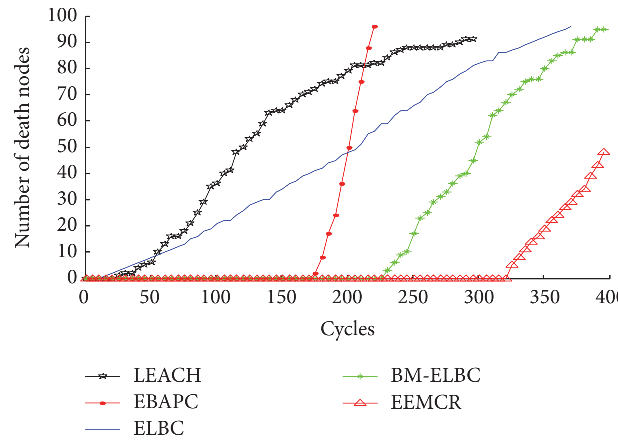


Figure 4: The death nodes versus the cycle of clustering.

Usually, the emergence of death nodes is later and number of death nodes is less, which indicates higher coverage preservation, less communication holes, better energy efficiency, and higher QoS. *Experiment 1* clearly shows the effect that the proposed EEMCR method has on the emergence of death nodes. The results of the experiment are presented in Figure 4. As shown in Figure 4, the emergence of death nodes with EEMCR is the latest one, and BM-ELBC algorithm is the second latest. In particular, the EEMCR method is far later than that of the other compared algorithms. When the cycle gets 400, almost half of the nodes are also surviving for the proposed EEMCR method; it is very prominent among the compared algorithms. The main reason for this situation is that EEMCR method examines an effective cluster head rotation mechanism, which effectively avoids the rapid energy consumption for the CHs, thereby slowing the death of nodes and enhancing the coverage preservation. On the other hand, since the EEMCR algorithm is based on the multihop transmission mode associated with the minimum aggregation tree-based generation methods, which builds node-joint paths with the smallest EC (e.g., Definition 2) between the CHs, that is, the optimum path, consequently, the energy consumption in the process of forwarding data to the base station significantly degrades. Namely, it improves the energy efficiency and reduces the monitoring blind area. Although the BM-ELBC algorithm also employs the multihop communication mode, it only considers the transmission capability rather than the energy losses on the multihop communication path according to the residual energy of the CHs. This case incurs fast energy consumption for the CHs with a higher opportunity of carrying out the data transmission tasks and leading to premature death of these CHs. In contrast, LEACH and ELBC algorithms firstly generate death nodes and the EBAPC algorithm is the second. This is firstly because LEACH, ELBC, and EBAPC algorithms are only based on single-hop communication mode. Additionally, the CHs election is relatively unreasonable in LEACH protocol.

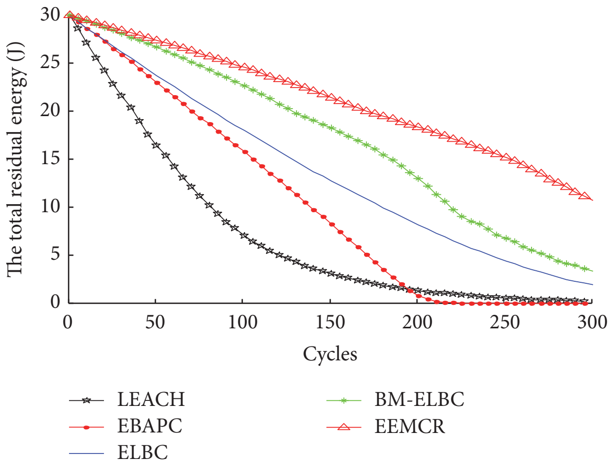


Figure 5: The residual energy versus the cycle of clustering.

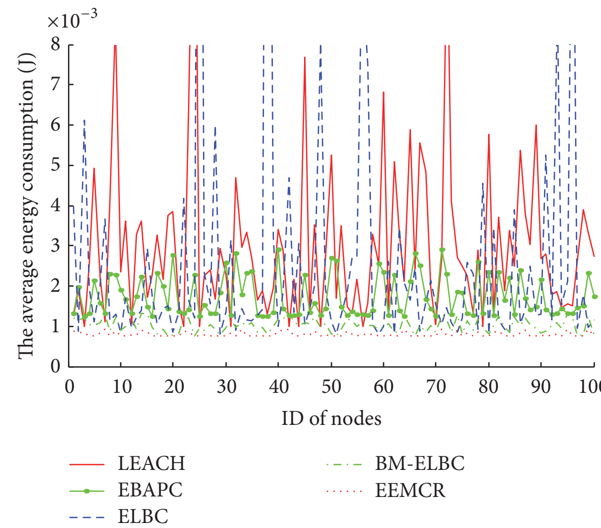
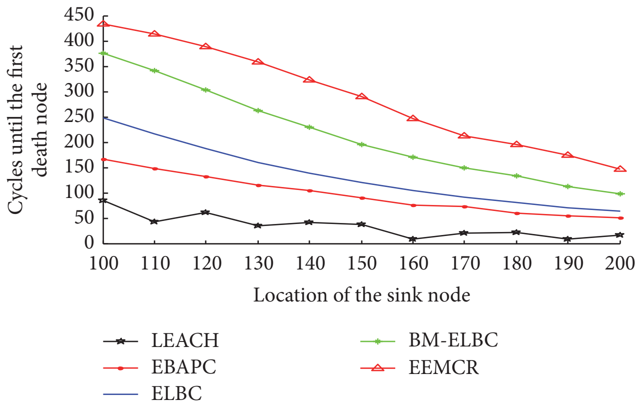


Figure 6: The average energy consumption of different nodes.

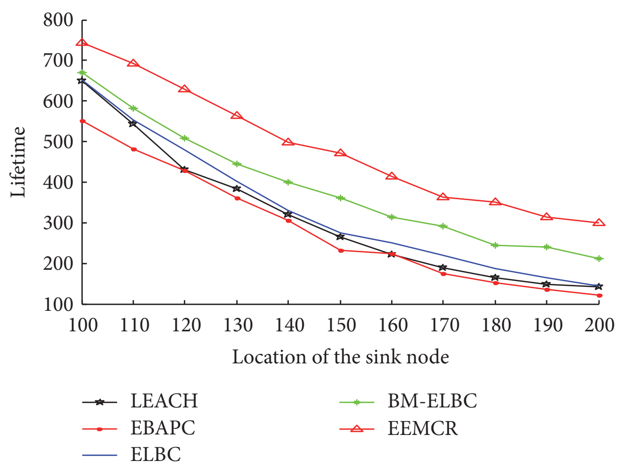
Additionally, EBAPC algorithm is not the algorithm whose emergence of death nodes is the earliest one, but its uptrend gradient is the largest one; that is, its ratio of death nodes is the highest. The primary reasons which caused this phenomenon are that the CHs have to undertake more data forwarding tasks after the cluster formation, which spur a fast energy consumption on some nodes; thus it is not even good as that of LEACH algorithm.

Experiment 2 demonstrates how the algorithms perform in relation to the speed of total energy consumption in the network. Generally, the speed of total energy consumption is one of the key indices to evaluate the comprehensive performance of networks. The more and faster the total energy consumption is, the shorter the survival time of nodes is. The experimental results are shown in Figure 5. It can be seen from Figure 5 that the descending trend of the total energy consumption of EEMCR method is the slowest; the BM-ELBC algorithm is the second slowest. They are superior to that of LEACH and EBAPC algorithm. This is because EEMCR and BM-ELBC algorithms use the multihop transmission approach, whereby leveraging the link redundancy to improve the transmission reliability and balance the energy consumption of CHs, as well as slowing down the speed of energy consumption for the whole network. Moreover, in comparison with the BM-ELBC algorithm, the EEMCR algorithm has outperformed it on saving energy. This is mainly because the optimum CH election and the selected clusters head always choose the best multihop routing for forwarding their aggregation data to the base station, thereby achieving an effective reduction of the insignificant energy consumption. In contrast, LEACH and EBAPC likely choose the nodes with less residual energy as CHs in the procedure of clustering, which causes fast energy consumption and leads some nodes to premature death.

Experiment 3 shows how the average energy consumption of nodes is affected by the algorithms. The experimental results are shown in Figure 6. Generally, the average energy consumption of nodes fully reflects equilibrium level of energy consumption, and also shows the optimization level of the selected multihop routing for data forwarding from another aspect. Under the same premise, the lower the average energy consumption of nodes is, the longer the node survives, the stronger the coverage preservation is, and the more reasonable the forwarding routing selection is. The closer the average energy consumption of each node, the more similar the energy consumption of different nodes, suggesting the self-organization and multihop transmission path of each node in WSNs are more reasonable and the energy consumption is more balanced, which may lead to a closer survival time of each node and more integrity of the network connectivity. As shown in Figure 6, the average energy consumption of nodes with the EEMCR algorithm is the most stable and is apparently lower than that of other compared algorithms. The BM-ELBC algorithm shows the second best. These experimental results indicate that the transmission paths using EEMCR algorithm are perfect, the energy consumption among different nodes achieves balancing, the survival time of the nodes is close to each other, and the capability to guarantee transmission reliability in WSNs is relatively strong. This is due to the EEMCR algorithm employing the evidence-efficient cluster head rotation mechanism to examine the clustering procedure, which results in a rational distribution of clusters and a proper size of clusters, as well as a balanced energy consumption among different nodes. Besides this, the selected minimum aggregation tree-based multihop routing using EEMCR algorithm leverages the link redundancy to improve transmission reliability and energy efficiency. Although the EBAPC algorithm also shows relatively stable average energy consumption among different nodes, its fluctuation is stronger than that of EEMCR algorithm and weaker than that of the remaining compared algorithms. The spurred reason is that the EBAPC algorithm also reasonably assigns the data forwarding tasks and disperses the energy consumption in WSNs to a certain extent. In contrast, LEACH and ELBC algorithms show a strong fluctuation of the average energy consumption on some nodes. For example, with the combined analysis shown in Figure 3(b),



(a) The emergence of the first death node versus the changing location of the sink



(b) The survival time versus the changing location of the sink node

Figure 7: The influence of base station location on the survival time.

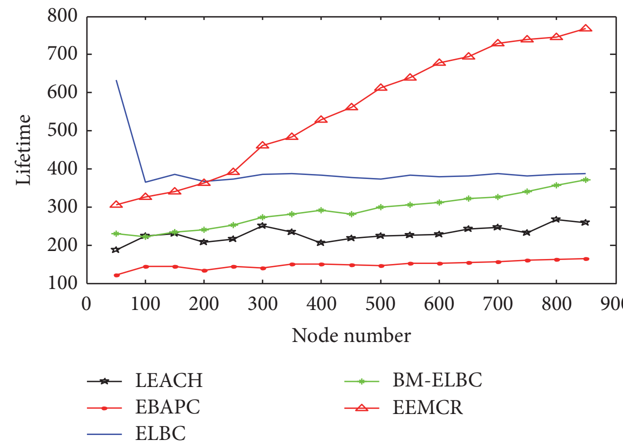


Figure 8: The influence of the scale on algorithms.

since the size of clusters is unreasonable using LEACH algorithm, some CHs have to undertake very large amounts of data transmission which results in the distinguished fluctuation in energy consumption. On the other hand, the average energy consumption between different nodes is very large, which indicates that the energy consumption is centralized on a few nodes and the energy consumption is unbalanced; the distribution of nodes shown in Figure 3(b) also verifies this point. Like that of LEACH algorithm, the similar fluctuation also happens in the ELBC algorithm; the primary reasons are due to the unreasonable CH election and lead to a tremendous variation of energy consumption for some local nodes. In summary, Figure 6 shows that the generated multihop communication path using EEMCR algorithm is reliable; data forwarding efficiency is relatively high, thus not only slowing the overall energy consumption, but also globally balancing energy consumption of different nodes.

Experiment 4 shows how changing the base station location affects the algorithm performance. Changing the location of the base station can better evaluate the robustness of the algorithm to the variation environment. The experimental results are shown in Figure 7. The base station moves from (100, 50) to (200, 50), namely, taking shifting horizontal coordinate as an example. The distance between the base station and the target area gets farther and farther, the emergence of the first death node with the five compared algorithms is getting more and more early, and the lifetime gets shorter and shorter. With respect to the descending trend, the EEMCR algorithm is the slowest which indicates that the emergence of the first death node is effectively postponed. At the same time, this phenomenon indicates that EEMCR algorithm has a better robustness to adapt to the changing configuration; this is mainly because it employs the efficient cluster head rotation mechanism while choosing a best multihop routing. And more, since considering the residual energy of nodes, it degrades the possibility of the CHs with lower residual energy as the intermediate forwarding nodes and significantly avoids their rapid energy consumption and even causing the premature death of the nodes undertaking forwarding tasks. Additionally, when the base station is farther, the BM-ELBC algorithm also delays the occurrence of the first death node, but it is still earlier than that of EEMCR algorithm; this is mainly because the BM-ELBC algorithm failed to choose the optimum communication path, which easily leads to increasing the energy consumption and shortening the network lifetime.

Experiment 5 demonstrates the algorithm capability to adapt to large-scale WSNs and the nodes changes between the ranges of 50–850. The experimental results are shown in Figure 8, in which the lifetime is taken as the primary evaluation index.

Usually, as the scale of the network increases, that is, the number of nodes in the network becomes larger, the workload also increases accordingly, and thus the CH election gets more critical. Figure 8 shows the network lifetime changes with the increasing number of nodes from 50 to 850. Especially after the number of nodes is greater than 200, the lifetime of the proposed EEMCR scheme is more prominent than that of the others, and the survival time of the network increases greatly. This situation indicates that the proposed EEMCR algorithm can effectively prevent the occurrence of premature death of nodes and prolong the network lifetime.

Namely, it indicates that it is suitable for the large-scale WSNs. Furthermore, the deeper reason is that taking the accumulated evidence of responsibility and availability as the deterministic criterion for the cluster head election is reliable and promising. On the other hand, the scheme of Kruskal-based construction of the minimum aggregation tree ensures that the generated communication path is optimum, which is effectively against link losses.

Additionally, besides the LEACH algorithm has a little fluctuation and ELBC has an abrupt change, the lifetime of the network generated by the remainder three algorithms shows more or less uptrend. The reason causing the changings in LEACH algorithm is the unreasonable CHs election mechanism. For ELBC algorithm, it has a relatively longer lifetime when the number of nodes is less than 200, its lifetime even reaches 621 within the size of 50 of nodes, and then the network lifetime performs stably and remains at around 390, which indicates that the ELBC algorithm only has a better adaptability to the small-scale network. The reason covered is primarily that the ELBC algorithm delivers data directly to the base station after the cluster formation, which makes it easy to encounter the bottleneck problem of data transmission such as data collision.

Based on the above experimental analyses and comparisons, the effectiveness of the proposed EEMCR scheme is convincing, and its efficiency is promising.

7. Conclusion and Future Works

In the light of network lifetime, coverage preservation, transmission reliability, and energy consumption, a novel EEMCR scheme has been proposed. In EEMCR scheme, the guaranteeing capability of a link or a node of data transmission is comprehensively considered. The support of nodes themselves and node-joint links to data transmission are abstracted as a propagation of responsibility and availability, that is, the accumulated evidence. The responsibility and availability integrate considerations on the residual energy of nodes, CH location, energy consumption on the selected communication path and distance between nodes. The presented cluster head rotation mechanism makes a globally reasonable CH distribution, proper size of clustering, and uniform allocation of energy consumption. On the other hand, backbone construction algorithm is employed to build the minimum aggregation tree within the candidate CHs, and the generated minimum aggregation tree provides an energy-efficient multihop routing to guarantee the optimal and shortest path between the CHs to the base station. This scheme can effectively improve coverage preservation, save energy, and even degrade the link redundancy. Although the EEMCR scheme inherits the infrastructure of LEACH algorithm, the theoretical analysis and empirical results demonstrate the promising performance over the LEACH algorithms. Regarding emergence of the first death node, equilibrium of energy assignment, rationality of CH distribution, and convergence rate of the proposed algorithm, our scheme is promising. The proposed scheme is applicable for the large-scale WSNs. In the future works, the performance of our scheme could be evaluated when embedded in the design of new routing protocols. Beyond that, various kinds of protocols design and topology construction methods based on the presented innovations for different layer of WSNs would be studied.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work is supported by the Future Research Projects Funds for the Science and Technology Department of Jiangsu Province (Grant no. BY2013015-23) and the Fundamental Research Funds for the Ministry of Education (Grant no. JUSRP211A 41).

References

[1] O. Durmaz Incel, A. Ghosh, B. Krishnamachari, and K. Chintalapudi, “Fast data collection in tree-based wireless sensor networks,” *IEEE Transactions on Mobile Computing*, vol. 11, no. 1, pp. 86–99, 2012.

[2] O. Gnawali, R. Fonseca, K. Jamieson, M. Kazandjieva, D. Moss, and P. Levis, “CTP: an efficient, robust, and reliable collection tree protocol for wireless sensor networks,” *ACM Transactions on Sensor Networks*, vol. 10, no. 1, article 16, 2013.

[3] Y. Jin, L. Wang, Y. Kim, and X. Yang, “EEMC: An energy-efficient multi-level clustering algorithm for large-scale wireless sensor networks,” *Computer Networks*, vol. 52, no. 3, pp. 542–562, 2008.

[4] A. P. Renold and S. Chandrakala, “Survey on state scheduling-based topology control in unattended wireless sensor networks,” *Computers and Electrical Engineering*, vol. 56, pp. 334–349, 2016.

[5] Z. X. Liu, Q. C. Zheng, L. Xue, and X. P. Guan, “A distributed energy-efficient clustering algorithm with improved coverage in wireless sensor networks,” *Future Generation Computer Systems*, vol. 28, no. 5, pp. 780–790, 2012.

[6] 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.

[7] C.-H. Lung and C. Zhou, “Using hierarchical agglomerative clustering in wireless sensor networks: an energy-efficient and flexible approach,” *Ad Hoc Networks*, vol. 8, no. 3, pp. 328–344, 2010.

[8] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, “Energy-efficient communication protocol for wireless microsensor networks,” in *Proceedings of the 33rd Annual Hawaii International Conference on System Siences (HICSS ’00)*, vol. 2, IEEE, January 2000.

[9] M. Cobos, F. Antonacci, A. Alexandridis, A. Mouchtaris, and B. Lee, “A Survey of Sound Source Localization Methods in Wireless Acoustic Sensor Networks,” *Wireless Communications and Mobile Computing*, vol. 2017, pp. 1–24, 2017.

[10] O. Younis and S. Fahmy, “Distributed clustering in ad-hoc sensor networks: a hybrid, energy-efficient approach,” in *Proceedings of the 23rd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM ’04)*, vol. 1, pp. 629–640, March 2004.

[11] B. Wang, H. B. Lim, D. Ma, and D. Yang, "A coverage-aware clustering protocol for wireless sensor networks," in *Proceedings of the 2010 6th International Conference on Mobile Ad-hoc and Sensor Networks, MSN 2010*, pp. 85–90, China, December 2010.

[12] L. M. C. Arboleda and N. Nasser, "Comparison of clustering algorithms and protocols for wireless sensor networks," in *Proceedings of the IEEE Canadian Conference on Electrical and Computer Engineering (CCECE '06)*, pp. 1787–1792, Ottawa, Canada, May 2006.

[13] C. Li, M. Ye, G. Chen, and J. Wu, "An energy-efficient unequal clustering mechanism for wireless sensor networks," in *Proceedings of the 2nd IEEE International Conference on Mobile Ad-hoc and Sensor Systems (MASS '05)*, pp. 597–604, Washington, DC, USA, November 2005.

[14] E. Leão, C. Montez, R. Moraes, P. Portugal, and F. Vasques, "Alternative Path Communication in Wide-Scale Cluster-Tree Wireless Sensor Networks Using Inactive Periods," *Sensors*, vol. 17, no. 5, p. 1049, 2017.

[15] L.-J. Chen, M. Liu, D.-X. Chen, and L. Xie, "Topology evolution of wireless sensor networks among cluster heads by random walkers," *Jisuanji Xuebao/Chinese Journal of Computers*, vol. 32, no. 1, pp. 69–76, 2009.

[16] Y.-H. Zhu, W.-D. Wu, V. C. M. Leung, and L.-H. Yang, "Energy-efficient tree-based message ferrying routing schemes for wireless sensor networks," in *Proceedings of the 3rd International Conference on Communications and Networking in China, ChinaCom 2008*, pp. 843–847, China, August 2008.

[17] Y. Jiang and H. Zhang, "Base station controlled intelligent clustering routing in wireless sensor networks," in *Advances in Artificial Intelligence*, vol. 6657 of *Lecture Notes in Comput. Sci.*, pp. 210–215, Springer, Heidelberg, 2011.

[18] K. X. Cui and Z. H. Li, "Clutering Network Topology Control Algorithm for EBAPC Based on Energy," *Computer Engineering*, vol. 38, no. 23, pp. 104–108, 2012.

[19] P. F. Li, "Energy-level Based Clustering Network Control Algorithm," *Computer Science*, vol. 41, no. 3, pp. 96–99, 2014.

[20] X. Yin, "Affinity propagation clustering algorithm based on node competing strength," *Computer Engineering and Applications*, vol. 51, no. 8, pp. 79–84, 2015.

[21] 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.

[22] Y. Tao, Y. Zhang, and Y. Ji, "Flow-balanced routing for multihop clustered wireless sensor networks," *Ad Hoc Networks*, vol. 11, no. 1, pp. 541–554, 2013.

[23] M. S. Levin and A. A. Zamkovoy, "Multicriteria Steiner tree with the cost of Steiner vertices," *Journal of Communications Technology and Electronics*, vol. 56, no. 12, pp. 1527–1542, 2011.

[24] B. J. Frey and D. Dueck, "Clustering by passing messages between data points," *American Association for the Advancement of Science: Science*, vol. 315, no. 5814, pp. 972–976, 2007.

[25] M. Younis, P. Munshi, G. Gupta, and S. M. Elsharkawy, "On efficient clustering of wireless sensor networks," in *Proceedings of the DSSNS 2006: 2nd IEEE Workshop on Dependability and Security in Sensor Networks and Systems*, pp. 78–87, USA, April 2006.

[26] O. Younis, M. Krunz, and S. Ramasubramanian, "Node clustering in wireless sensor networks: recent developments and deployment challenges," *IEEE Network*, vol. 20, no. 3, pp. 20–25, 2006.

[27] W. Zhou, H.-M. Chen, and X.-F. Zhang, "An energy efficient strong head clustering algorithm for wireless sensor networks," in *Proceedings of the 2007 International Conference on Wireless Communications, Networking and Mobile Computing, WiCOM 2007*, pp. 2584–2587, China, September 2007.

[28] A. Chamam and S. Pierre, "A distributed energy-efficient clustering protocol for wireless sensor networks," *Computers and Electrical Engineering*, vol. 36, no. 2, pp. 303–312, 2010.

[29] M. Khalily-Dermany, M. Shamsi, and M. J. Nadjafi-Arani, "A convex optimization model for topology control in network-coding-based-wireless-sensor networks," *Ad Hoc Networks*, vol. 59, pp. 1–11, 2017.

[30] B. Yin, H. Shi, and Y. Shang, "An efficient algorithm for constructing a connected dominating set in mobile ad hoc networks," *Journal of Parallel and Distributed Computing*, vol. 71, no. 1, pp. 27–39, 2011.

[31] S. Lindsey and C. S. Raghavendra, "PEGASIS: power-efficient gathering in sensor information systems," in *Proceedings of the IEEE Aerospace Conference*, vol. 3, pp. 1125–1130, Big Sky, Mont, USA, March 2002.

[32] L. Keller, E. Atsan, K. Argyraki, and C. Fragouli, "SenseCode: network coding for reliable sensor networks," *ACM Transactions on Sensor Networks*, vol. 9, no. 2, article 25, 2013.

[33] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W. H. Freeman, San Francisco, Calif, USA, 1979.