A Collaborative Data Gathering Mechanism Based on Fuzzy Decision for Wireless Sensor Networks

Wen Si, Hui-yuan Shi, and Pan Liu

1College of Information and Computer Science, Shanghai Business School, Shanghai 201400, China
2Department of Environmental Science and Engineering, Fudan University, Shanghai 200433, China
3Shanghai Key Laboratory of Computer Software Testing & Evaluating, Shanghai 200235, China

Correspondence should be addressed to Pan Liu; liup@sbs.edu.cn

Received 27 May 2014; Revised 26 July 2014; Accepted 1 August 2014

1. Introduction

In recent years, wireless sensor networks (WSNs) have seen rapid development in many applications, such as in environmental monitoring, target tracking, outer space exploring, and industrial automation [1]. In spite of the increasing demand for all kinds of sensing services and applications, there still lacks a clear understanding of collaborative techniques to design collaborative protocols for wireless sensor networks. This paper proposes collaborative data gathering mechanism based on fuzzy decision for wireless sensor networks. The proposed algorithm integrates some key parameters, for example, nodes’ residual energy level, the number of neighbors, centrality degree, and distance to the sink, into fuzzy decision. Numerical and simulation results validate the proposed algorithm for the networks in finding the optimum cluster heads and realizing better performances in clusters distribution and energy efficiency improvement.

Routing and data gathering are the major considerations in designing the pattern and operation modes of the WSNs. A routing algorithm achieving significant energy efficiency is cluster routing [4]. At present, the cluster organization is widely used in wireless sensor networks. The advantages of using clustering techniques for data collection are as follows: (1) cluster head (CH) node is responsible for long-distance transmission and forwarding of data to ensure the original data communications coverage area; (2) the cluster members can close the communication unit whenever it is necessary, so that the network’s energy consumption can be effectively saved; (3) clustering topology is simple and easy to manage and is conducive to the application of distributed algorithms; (4) the protocol for communication between nodes is adaptive and just relies totally on the neighbor node information to decide whether to join or become a cluster head, so that network routing control information can be reduced; (5) the data from all members gathered in the cluster should be aggregated by cluster head, and it can greatly reduce the amount of network traffic.

This paper considered the remaining energy level of sensor nodes, centrality degree, and the number of neighbor nodes and consequently determined the weight of each
attribute for cluster head selection by the fuzzy decision, which can effectively increase the high level of residual energy of sensor nodes become cluster head node probabilities and ensure that the distribution of cluster heads is more uniform and reasonable.

For reducing the communication overhead, data aggregation method was presented in wireless sensor networks. For the sensor network with high nodes density, there may be a large amount of redundant data from those who detect the objects in the same domain [5]. According to the relevance of sensing data, error function and fuzzy correlation function can be designed to acquire a comprehensive support degree among the nodes, and then it is possible to determine the reliability of each sensor and improve precision of the data fusion result. As far as the accuracy of cluster head data gathering was concerned, it can effectively reduce the amount of intracluster data and improve the energy efficiency [6].

This paper specifically focuses on exploring energy-efficient data gathering mechanism for collaborative WSNs to make WSN-based applications more reliable and effective in industry-related scenarios. In this paper, a collaborative data gathering mechanism based on fuzzy decision (DGM-FD) is proposed, aiming at distributing the load among all sensor nodes evenly. In order to reach this objective, some critical parameters, such as nodes’ residual energy level, the number of neighbors, centrality degree, and distance to the sink, are taken into account during the cluster head competition. Numerical and simulation results validate that the proposed algorithm can effectively find the optimum cluster heads and realize performances both in proper distribution of clusters and in improving the energy efficiency of the networks.

The specific contributions of this paper include the following:

(i) a literature survey about various existing data gathering algorithms and analysis of their advantages and disadvantages,
(ii) an effective collaborative data gathering mechanism based on fuzzy decision for wireless sensor networks is proposed,
(iii) an algorithm (DGM-FD) for collaborative data gathering in wireless sensor networks is proposed,
(iv) performance analysis of the proposed algorithm and an evaluation of the algorithm with respect to other existing algorithms.

The rest of the paper is organized as follows. We present the related works in Section 2. The energy-efficient data gathering mechanism is formally discussed in Section 3. The proposed DGM-FD model is introduced in detail in Section 4. In Section 5 the proposed mechanism is evaluated and finally we conclude in Section 6.

2. Related Works

In this section, we present a review of the recent developments of collaborative wireless sensor networks. The focus of attention varies from application specific detection to enhancement of middleware. In order to gather information more efficiently in energy consumption, clustering algorithm is introduced into the applications of wireless sensor networks. In WSNs, nodes can be partitioned into a number of small groups called clusters. Each cluster has a coordinator, referred to as a cluster head, and a number of member nodes. The cluster heads are responsible for aggregating the collected data and forwarding it to the base station (BS) through other cluster heads in the network. By rotating cluster heads periodically, the energy consumption of the sensor nodes over the network can be balanced.

Several WSNs applications require only an aggregate value be reported to the observer. In this case, sensors in different regions of the field can collaborate to aggregate their data and provide more accurate reports about their local regions [7]. Data aggregation reduces the communication overhead in the network, leading to significant energy savings.

Meghanathan proposed two distributed algorithms to construct (i) stable predicted link expiration time-based data gathering (LET-DG) trees that also incur lower delay per round as well as larger throughput per tree and (ii) energy-efficient minimum-distance spanning tree based data gathering (MST-DG) trees that incur larger node and network lifetimes and inflict lower coverage loss on the underlying network at any time instant [8]. Bober and Bleakley proposed BailighPulse, an energy efficient data gathering protocol for mostly-off WSN applications [9]. BailighPulse incorporates a novel multihop wake-up scheme that allows for energy efficient recovery of network synchronization after long off periods. Ebrahim and Assi proposed MSTP [10], a new method for data aggregation in large-scale WSN using compressive sensing and random projection. The proposed method selects random projection nodes to generate routing trees with each projection node gathering a weighted sum from all the nodes in the network. Gupta et al. presented the problem of static itinerary based Agent migration protocol in WSN [11]. ETMAM (energy and trust aware mobile agent migration) presented an integrated solution for reliable agent migration within network. In ETMAM, energy and trust are both considered for deciding the next node for the agent migration.

EECS algorithm had improved LEACH (low-energy adaptive clustering hierarchy) method by changing probability. In probability function [12], energy parameter has been considered to choose cluster heads. Also reduction in search space has increased clustering speed. Li et al. proposed an unequal clustering algorithm (EEUC), where the clusters close to the base station have smaller sizes than clusters far from the base station [13]. However, it may produce lone nodes since the cluster head election is probabilistic.

Fuzzy logic showed its ability to cope with information with a high degree of uncertainty in heterogeneous engineering fields [14]. For some problems, it is usually assumed that there is fuzziness in each objective due to the imprecise nature of judgment as a decision maker [15]. The fuzzy mechanism can be used to find a compromised solution, which looks at the way the solutions are contributing to each objective and assigns a fuzzy variable [16]. The fuzzy mechanism supplies
3. Collaborative Data Gathering Mechanism

3.1. Collaborative Data Gathering. At present, in most of the existing clustering algorithms, a certain probability of nodes is chosen to become cluster heads that is usually based on the residual energy of nodes, the average energy, neighbors, and other factors [23]. In case of an uneven distribution of network, the energy consumption of nodes is likely to cause uneven and rapid death and is resulting in the coverage holes problem. Therefore, while designing clustering data collection protocol for WSNs, we should consider the following aspects: firstly, clustering algorithm uses fully distributed control mechanisms and overcomes the shortcomings of inadaptability in centralized way. Secondly, compared to the direct routing, clustering algorithm can uniformly manage the distributed nodes in the network so as to achieve the load balancing of the whole network. Thirdly, for avoiding excessive energy consumption of certain nodes with high burden, the distribution of cluster heads should be chosen as evenly as possible.

The collaborative strategy ensures the efficiency and the robustness of the data gathering, while limiting the required communication bandwidth. Data aggregation has emerged as a useful paradigm in sensor networks. The key idea is to combine data from different sensors to eliminate redundant transmissions and provide a rich, multiperspective view of the environment being monitored. However, most of the research works focus on reducing the energy consumed by the sensors during the process of data gathering, such as finding routes between pairs of end nodes. On the other hand, some researches focus on the intrinsic characteristics of the collected data. In some scenarios, sensors are deployed to monitor continuous environmental conditions such as temperature, humidity, or seismic activity and periodically produce relevant information by sensing an extended geographic area that is eventually transmitted to the sink for processing. Since different sensors partially monitor the same spatial region result in the correlated data, the data aggregation can reduce the energy consumption in transmission.

For the organization model of clustering, due to the members distributed in an immediate area, there is a strong correlation of the data collected in the cluster.

In the following sections, we will analyze and define the critical elements for cluster head selection, and a correlation function in fuzzy theory is adopted for the intracluster data aggregation.

3.2. Centrality Degree. Due to the uneven distribution of nodes in wireless sensor network, the centrality degree of cluster head will influence the whole intracluster energy dissipation. The more the cluster head deviated from the geometric centre of the cluster, the more energy the member node will consume for transmitting the message. Assume that node $i$ is the candidate cluster head with a coordinate which is $(x_i, y_i)$. The positions of the all nodes in the range of node $i$ can be termed as $(x_1, y_1), (x_2, y_2), \ldots, (x_m, y_m)$, respectively. In an ideal condition, we can give the energy function for the intracluster communication of cluster head $i$:

$$E_{\text{bread}} = \sum_{j=1}^{m} \gamma \left[ (x_i - x_j)^2 + (y_i - y_j)^2 \right], \quad (1)$$

where $\gamma$ is constant coefficient. It can be derived from (10), and it can be observed that $E_{\text{total}}$ is the most optimal on the condition that the cluster head $i$ located in \((1/m) \sum_{j=1}^{n} x_j, (1/m) \sum_{j=1}^{n} y_j\).

The centrality degree of cluster head $i$ is defined as follows:

$$\text{Centri}_i = \frac{\sqrt{(x_i - (1/m) \sum_{j=1}^{m} x_j)^2 + (y_i - (1/m) \sum_{j=1}^{m} y_j)^2}}{r_0}, \quad (2)$$

where $r_0$ is the defuzzified output of fuzzy if-then rules. The tentative cluster head becomes a cluster head if it has the greatest chance in its vicinity. There are distributed and centralized fuzzy logic clustering approaches. A fuzzy energy-aware unequal clustering algorithm (EAUCF) was proposed to address the hot spots problem [18]. EAUCF aimed to decrease the intracluster work of the cluster heads that either are close to the base station or have low remaining battery power. A fuzzy logic approach was adopted in order to handle uncertainties in cluster head radius estimation. The Gupta fuzzy protocol using the fuzzy logic approach to select CHs utilizes three parameters: energy level, concentration, and centrality [19]. The protocol is used to collect energy level and location information for each node in the setup stage.

The CHEFL (cluster head election mechanism using fuzzy logic) protocol used a fuzzy logic approach to maximize the lifetime of WSNs [20]. It was similar to the Gupta protocol but it does not need the BS to collect information from all nodes. Instead, the CHEFL protocol uses a localized CH selection mechanism with fuzzy logic. The LEACH-FL (LEACH protocol using fuzzy logic) protocol was proposed in [21]. This protocol used fuzzy logic to improve the LEACH protocol by considering three different parameters: energy level, node density, and distance between the CH and the BS. This model was the same as the Gupta protocol in the sense of a setup stage and a steady-state stage. In order to choose cluster heads, a two-level fuzzy method that included local level and global level was used [22]. In local level, node’s capability of being cluster head can be evaluated based on two parameters: energy and the number of neighbors. In global level, three parameters had been considered: centrality degree, closeness to base station, and the distance between cluster heads.
where $r_0$ is the communication radius of the sensor and $m$ is the number of member nodes in the cluster. With the consideration of different cluster size, the centrality degree can be derived as follows:

$$f_1(i) = \text{Centri}\_\text{degree}(i) = \ln m \times \left(1 - \left(\left(\frac{x_i - 1}{m} \sum_{j=1}^{m} x_j\right)^2 + \left(\frac{y_i - 1}{m} \sum_{j=1}^{m} y_j\right)^2\right)^{1/2} r_0^{-1}\right).$$

Cluster head with high centrality degree can help reduce intracluster communication energy. Therefore, the centrality degree of candidate cluster head can contribute to the rational distribution of cluster heads.

### 3.3. The Number of Neighbor Nodes

In this paper, the network coverage ratio $\text{COV}_n$ is defined as the mathematical expectation of coverage ratio of the whole monitored area:

$$\text{COV}_n = 1 - (1 - p)^n,$$

where $p$ is the probability of coverage monitoring area for each sensor node and $n$ is the number of nodes in the monitored area. For node $i$, its sensing range $\text{SRange}_i$ is covered by $k$ neighbor nodes at least. Then, we have

$$\text{COV}_k(i) = 1 - (1 - p)^k,$$

$$\rho = \frac{\sum_{j=1}^{\text{set}(i)} \text{SRange}_j}{|\text{set}(i)| \times \text{SRange}_i},$$

where $\text{set}(i)$ is the set of neighbor nodes of node $i$ and $k$ is the number of all neighbor nodes in the sensing range. In practical applications, the coverage ratio of network only meets the corresponding specified coverage ratio. Assume that $\delta$ is the lower limit of the specified coverage; we can get the number of neighbor nodes that meet the coverage ratio of node $i$ in terms of sensing range $\text{SRange}_i$; that is to say, $\text{COV}_k(i) = \delta$:

$$\text{Neighb}(i) = \frac{\ln(1 - \delta)}{\ln(1 - p^2)}.$$

The optimum number of clusters $k_{\text{opt}}$ for the $n$ sensor nodes distributed randomly in a $(M \times M)$ field and $k_{\text{opt}}$ is defined as follows [24]:

$$k_{\text{opt}} = \sqrt{\frac{n}{2\pi}} \times \sqrt{\frac{\varepsilon_{fs}}{\varepsilon_{mp}}} \times \frac{M}{d^2}.$$

Hence, the average number of neighbors is $N/k_{\text{opt}}$ for a cluster. And the deviations between the neighbors of a candidate cluster head and the average number can be calculated by the following equations:

$$f_2(i) = \frac{\text{Neighb}(i) - \frac{N}{k_{\text{opt}}}}{\ln(1 - \delta)} = \frac{\ln(1 - \delta)}{\ln(1 - p^2)} - \frac{M^2 \sqrt{2\pi\varepsilon_{mp}}}{d^2 \sqrt{\varepsilon_{gc}}}.$$

### 3.4. Energy Level

In the cluster head election, the remaining energy of nodes is a very important factor. The selected nodes which are with plenitudinous residual energy and are relatively close to the sink can well extend the lifetime of network. In order to optimize the energy consumption, this paper adopts the concept of residual energy level.

Assume that initial energy of nodes is $E_0$, and the initial energy of the node is divided into grades $1 \sim p$. In the initial phase, most of the nodes are of abundant energy, and the probability of premature death of node is low. In contrast, in the final stage, the residual energy of node is relatively small. Therefore, we define the energy of each level according to the following formula:

$$\text{REL}_i = \left(\frac{(i - 1)^2}{\rho^2}E_0, \frac{i^2}{\rho^2}E_0\right).$$

With the energy dissipation of the sensors, the level of the residual energy will vary from 1 to $p$. When the remaining energy is reduced, the energy differential is smaller. This indicates that the gap of energy level is getting smaller while the node declines. During the selection of cluster heads, the nodes with high energy level have an advantage. Thus, the energy level of nodes should reasonably reflect the nodes’ residual energy:

$$f_3(i) = \frac{\ln(1 + \theta \text{REL}_i)}{\ln(1 + \theta)},$$

where $\text{REL}_i$ is the energy level of node $i$ and $\alpha$ is the adjustable parameters.

### 3.5. Distance

According to the radio energy dissipate model, the node closer to the sink is elected as a cluster head which consumes the lower power than the one far away. Therefore, the distance function is defined as follows:

$$f_4(i) = 1 - \frac{\text{DS}_i}{\text{DS}_{\text{max}}} = 1 - \frac{\max_{i \leq j \leq n} \left\{\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}\right\}}{\sqrt{(x_i - x_0)^2 + (y_i - y_0)^2}}.$$

where $\text{DS}_i$ is the distance from node $i$ to the base station and $\text{DS}_{\text{max}}$ is the node most distant to the base station in WSNs.

The four indicators described above for the reasonable selection of cluster head have a very important impact. In the next section, we propose the method based on the correlation functions and fuzzy nearness to analyze the proportion of these critical indicators in the cluster heads selection. It will effectively avoid the situation where the lower residual energy node becomes a cluster head node or the distribution of cluster head is uneven.
4. The Proposed DGM-FD Mechanism

The most important thing is to select appropriate cluster heads according to the above indicators for balancing energy expenditure and improving the network lifecycle. Fuzzy decision can be an effective method to sort the complex objectives in the domain of decision theory or select the optimal target resolution with respect to the restriction of fuzzy logic.

The most important factors for the nodes competing for being cluster head (i.e., the remaining energy level of sensor nodes, centrality degree, and the number of neighbor nodes) are fully considered. It would be simplistic and subjective if it is to just use the weighted sum of the evaluation results. Therefore, we adopt the method of fuzzy mathematics to obtain the comprehensive evaluation, and the objective weights of attributes are determined through the entropies according to the positive relationship between the values of the samples. According to the sampled index value of all nodes in the cluster, the ideal solution and the negative ideal solution are calculated. Next, the distance between the object nodes in the cluster, the ideal solution and the negative ideal solutions in (14),

\[
\begin{align*}
\text{ideal solution} & = (u_1^*, u_2^*, \ldots, u_n^*) \\
\text{negative ideal solution} & = (u_1^-, u_2^-, \ldots, u_n^-)
\end{align*}
\]

where \( u_k^+ \) is the distance from the ideal solution and \( u_k^- \) is the distance from the negative ideal solution and both are defined as follows:

\[
\begin{align*}
\text{distance} & = \max \{u_1, u_2, \ldots, u_n\} \\
\text{negative distance} & = \min \{u_1, u_2, \ldots, u_n\}
\end{align*}
\]

According to the information theory [25], the relative entropy of systems \( \alpha \) and \( \beta \) is defined as follows:

\[
\text{Corr} = \sum_{i=1}^{N} \left( \alpha_i \ln \frac{\alpha_i}{\beta_i} + (1 - \alpha_i) \ln \frac{1 - \alpha_i}{1 - \beta_i} \right)
\]

In terms of the definition of the relative entropy, \( \lambda_k^+ \) denotes the relative entropy is associated with attribute \( k \) and ideal solution \( u_k^+ \), and \( \lambda_k^- \) is the relative entropy associated with attribute \( k \) and positive ideal solution \( u_k^- \). There exist the following expressions:

\[
\begin{align*}
\lambda_k^+ & = \sum_{j=1}^{m} \left( \lambda_k^+ \ln \frac{u_k^+}{u_{jk}} + (1 - \lambda_k^+) \ln \frac{1 - u_k^+}{1 - u_{jk}} \right), \\
\lambda_k^- & = \sum_{j=1}^{m} \left( \lambda_k^- \ln \frac{u_k^-}{u_{jk}} + (1 - \lambda_k^-) \ln \frac{1 - u_k^-}{1 - u_{jk}} \right)
\end{align*}
\]

The relative closeness to the ideal solution is calculated by the following equation: \( \text{Corr}_k = \lambda_k^+ / (\lambda_k^+ + \lambda_k^-) \). According to the relative closeness of all attributes value to the ideal solution or the negative ideal solution, the vector of the weight in the decision model can be obtained as follows:

\[
\omega_k = \frac{1 - \text{Corr}_k}{\sum_{k=1}^{l} (1 - \text{Corr}_k)} \left( \lambda_k^+ / (\lambda_k^+ + \lambda_k^-) \right)
\]

Finally, \( k_{\text{opt}} \) nodes can be chosen as a relatively large value, which will be selected as cluster head in the next round.

In densely distributed network, there is an abundance of information that can be collected by sensors. For the nodes especially in one cluster, they usually locate in an adjacent region and monitor same objects. In order to minimize the volume of the transmitted data, we can design an aggregation scheme to exploit spatiotemporal correlations in the readings, which is obtained by the nodes in the network. When the target object is detected, the readings from different sensing nodes often exhibit certain correlation. Obviously, it is possible that there is a lot of redundant data which has a strong correlation in the process of perception. With some nodes being scheduled into the sleep state to avoid the production of these redundant data, it will undoubtedly reduce network energy consumption and prolong the network lifetime.

Therefore, the correlation function in fuzzy theory is presented to calculate the mutual support degree between the nodes during the process of the data aggregation in a single cluster. Then the redundant nodes, which should be changed into sleeping mode, can be chosen in order to save the energy. After receiving the sensed data of all members, the cluster heads will start analyzing perceptual characteristics of these readings. Moreover, as far as cluster density distribution was
concerned, the more the nodes with equal or approximate coverage area in the network, the faster the running speed and the more the residual energy. Generally speaking, the nodes’ sensibility of data has a positive correlation with the comprehensive support degree.

Assume that \( X_i \) and \( X_j \) are the perceptual data from nodes \( n_i \) and \( n_j \), respectively. With each node's sensing data subject to Gaussian distribution, the probability distribution functions can be described as \( p_i(x) \) and \( p_j(x) \). \( x_i \) and \( x_j \) denote the data collected in an interval: then the perception data confidence distance between \( n_i \) and \( n_j \) can be measured as follows:

\[
d_{ij} = 2 \int_{x_i}^{x_j} p_i \left( \frac{x}{x_j} \right) dx,
\]
\[
d_{ji} = 2 \int_{x_i}^{x_j} p_j \left( \frac{x}{x_j} \right) dx,
\]
\[
p_i \left( \frac{x}{x_j} \right) = \frac{1}{\sqrt{2\pi\sigma_i}} \exp \left\{ -\frac{1}{2} \left( \frac{x-x_i}{\sigma_i} \right)^2 \right\},
\]
\[
p_j \left( \frac{x}{x_j} \right) = \frac{1}{\sqrt{2\pi\sigma_j}} \exp \left\{ -\frac{1}{2} \left( \frac{x-x_j}{\sigma_j} \right)^2 \right\}.
\]

According to confidence distance measurement, confidence distance matrix sensor data \( D_n \) can be constructed as follows:

\[
D_n = \begin{bmatrix}
d_{11} & d_{12} & \cdots & d_{1n} \\
d_{21} & d_{22} & \cdots & d_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
d_{n1} & d_{n2} & \cdots & d_{nn}
\end{bmatrix} \tag{20}
\]

Next, we can obtain the supported degree of the data by the other nodes based on the method of fuzzy logic. In general, the threshold \( \xi \) is given, and a random \( d_{ij} \) the integration factor is calculated as follows:

\[
r_{ij} = \begin{cases} 0, & d_{ij} \leq \xi, \\ 1, & d_{ij} > \xi. \end{cases} \tag{21}
\]

If \( r_{ij} = 0 \), the relationship of the supported data between \( n_i \) and \( n_j \) is poor. Whereas \( r_{ij} = 1 \), the relationship is tightened. The problem is that the value of threshold \( \xi \) is too absolute and subjective. Thus, the actual situation is not conducive to make an objective judgment. For evaluating the effectiveness of the readings of all nodes, the correlation function of fuzzy theory is used to determine the supported degree for each node in the cluster. Suppose \( r_{ij} = 1 - d_{ij} \); the correlation function can be defined as follows:

\[
g(i \mid j) = \frac{r_{ij}}{\max \{r_{ij}, r_{ji}\}}. \tag{22}
\]

The threshold \( \xi(i) \) for node \( n_i \) can be calculated according to (15):

\[
\xi(i) = \min \{ g(i \mid j) \}, \quad j = 1, 2, \ldots, n. \tag{23}
\]

And the correlation matrix \( R(i) \) can be derived as follows:

\[
R(i) = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\
r_{21} & r_{22} & \cdots & r_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
r_{n1} & r_{n2} & \cdots & r_{nn}
\end{bmatrix}, \tag{24}
\]

where \( r_{ij} \) denotes the supported degree of \( n_i \) from \( n_j \), \( 0 \leq r_{ij} \leq 1 \).

The comprehensive level of support can be finally obtained as \( s(i) = 1/(1 + \|R(i)\|^2) \).

In order to minimize the energy consumption, the nodes with low supported degree can be filtered as the redundant nodes, which will be in sleep mode in the next rounds. The rounds of sleep \( T_{\text{sleep}} \) are defined by the following equation:

\[
n' = n \times \varphi, \tag{25}
\]
\[
T_{\text{sleep}} = \frac{E_0}{E_{\text{res}}(i)} \times T_{\text{active}},
\]

where \( \varphi \) is the percentage of redundant nodes, \( T_{\text{active}} \) is the working rounds of the cluster head, and \( E_0 \) is the initial energy of the node.

The number of rounds that the redundant nodes are scheduled to be in sleep mode is associated with the node residual energy.

5. Simulation Experiment

In this section, the experiments are implemented to evaluate DGM-FD. DGM-FD is compared with two different data gathering algorithms, namely, EECS [12] and FSC [26]. EECS and FSC are known as the classical methods for collaborative data gathering in wireless sensor networks. Several experiments are conducted on NS2 simulator to evaluate these algorithms. In order to evaluate the proposed algorithm, three different scenarios are developed. In the first scenario, the base station is located at the center of the WSNs. In the other two scenarios, the base station is outside the WSNs. In 100 \( \times \) 100 square area, the distribution of clusters in the network is compared to DGM-FD and FSC. Assume that the transmission radius of the sensor node is 20 meters and each cluster head forwards the aggregated data to the base station directly without using a relay node. In these experiments, the number of deployed sensors is 400 and we choose the formed clusters in the random round for analysis and comparison with different position of the sink node.

In order to extend the lifetime of sensor nodes, FSC makes use of fuzzy logic for cluster head selection. In this section, we firstly utilized FSC to examine the distribution of nodes in the WSNs. Figures 1(a) and 1(b) show the snapshots of proposed algorithm and FSC, respectively. When the sink node locates at the center of the network, there are 65 clusters in DGM-FD and the number of isolated nodes is 6. The average number of nodes in the cluster is 6.15, and the cluster size variance is 4.07. FCS produced 41 cluster heads and 5 isolated clusters. The average cluster size is 9.75 with its variance 7.31. This means that DGM-FD controls the size of each cluster better and
obtains less fluctuation in the number of nodes in all clusters. In other words, there is not much difference in the number of members for most clusters, and these cluster heads will not undertake excessive communication burden. In FCS, the scale of the clusters is much larger, and the cluster head will have to bear a heavier traffic load and sometimes may lead to premature death.

In the next scenario, the base station is located at the edge area of the wireless sensor network. The sink coordinates on this scenario are (100,100), and the topology of cluster formation is shown in Figure 2. DGM-FD produced 66 cluster heads and 5 isolated clusters, in comparison to 39 cluster heads and 5 isolated nodes. Nevertheless, the cluster size variance of FCS is 8.37, which is still much larger than 3.68 in DGM-FD.

In the third scenario, the base station locates at the edge area of the wireless sensor network and the position of sink is (100, 50), and the topology of cluster formation is shown in Figure 3. It is observed that there are 66 clusters in DGM-FD. Meanwhile, there are 7 isolated nodes with a cluster size variance 3.98. Comparatively, 41 clusters come into being in FCS, and the cluster size variance is 5.37. Due to the factors of centrality of cluster head taken into account, most cluster heads generally located in the center of the cluster, which can significantly reduce energy expenditure. Since the average size of the cluster in DGM-FD is smaller than that in FCS, the former can afford relative low burden.

Figure 4 shows the distribution of the alive sensor nodes with respect to the number of rounds for each simulated algorithm. As shown in this figure, DGM-FD is apparently the most energy-efficient algorithm. The sensor nodes of EECS and FSC start to die in the earlier rounds. The death time of the first node of DGM-FD is much smaller than that of nodes of all the other algorithms. DGM-FD provides at least
1970 stable rounds for these particular WSNs, whereas EECS provided 1080 rounds and FSC 1490 rounds. Furthermore, the energy of the network exhausts gradually about 2250 rounds at EECS and 2440 rounds at FSC. But DGM-FD cannot undertake 2700 rounds even when all nodes die. This figure clearly shows that the proposed algorithm is more stable than the other algorithms, in which sensor node deaths begin later for DGM-FD and continue linearly until all sensor nodes die.

The distribution of the average energy consumption of all nodes with respect to the number of rounds for each algorithm is depicted on a fast line chart in Figure 5. The simulation results show that the average energy consumption of all nodes varies comparatively stably during the most of the rounds in DGM-FD. The average energy consumption of other algorithms fluctuates in a wide range, especially in EECS. On the other hand, the average energy consumption of DGM-FD can keep in the low level, which benefits from the optimization of the selection of cluster heads.

During the formation of clusters in DGM-FD, the residual energy of nodes is paid full consideration, and the node with high energy level is selected as cluster head. The factors about the number of adjacent nodes and centrality degree are introduced, which makes the cluster size even and the cluster head locates near the centrality of its cluster as much as possible, while in EECS, due to the random selection of the cluster head, low residual energy nodes may also be selected.
as cluster head and in this case the death of nodes of low energy level would be accelerated.

The energy dissipation of all nodes in the three algorithms, as well as the variation tendency, is illustrated in Figure 6. As shown in the figure, DGM-FD consumes much less energy than FSC and EECS. As explained earlier, the optimal cluster heads selection is adopted at every round, and the reasonable intracluster data aggregation mechanism is added in DGM-FD, resulting in a significant energy-saving effect.

6. Conclusions

To improve the energy-efficiency and achieve the network load balancing, a DGM-FD mechanism is proposed in this paper. DGM-FD aims to distribute the workload among all sensor nodes evenly. Moreover, a correlation function in fuzzy theory is adopted for the intracluster data aggregation. The experimental results show that the proposed algorithm can effectively optimize the selection of cluster heads and obtain better performances both in proper distributing of cluster heads and in improving the energy efficiency of the networks. In the future, we will explore energy-efficient data gathering mechanism for collaborative WSNs to make WSN-based applications more reliable and effective in industry-related scenarios.

Conflict of Interests

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

Acknowledgments

This research work was supported by Shanghai Natural Science Fund (no. 14ZR1429800), Teachers Training Plan of Shanghai Municipality under Grant (nos. SXY12003 and ZZSD13008), and National Natural Science Foundation of China (nos. 61170044 and 61073050). And the authors wish to thank the anonymous reviewers who helped to improve the quality of the paper.

References

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, “Wireless sensor networks: a survey,” Computer Networks, vol. 38, no. 4, pp. 393–422, 2002.
[2] K. Akkaya and M. Younis, "A survey on routing protocols for wireless sensor networks," Ad Hoc Networks, vol. 3, no. 3, pp. 325–349, 2005.
[3] X. Zhu, L. Shen, and T.-S. P. Yum, “Hausdorff clustering and minimum energy routing for wireless sensor networks,” IEEE Transactions on Vehicular Technology, vol. 58, no. 2, pp. 990–997, 2009.
[4] N. Bouabdallah, M. E. Rivero-Angeles, and B. Sericola, "Continuous monitoring using event-driven reporting for cluster-based wireless sensor networks," IEEE Transactions on Vehicular Technology, vol. 58, no. 7, pp. 3460–3479, 2009.
[5] C.-M. Chao and T.-Y. Hsiao, “Design of structure-free and energy-balanced data aggregation in wireless sensor networks,” Journal of Network and Computer Applications, vol. 37, no. 1, pp. 229–239, 2014.
[6] R. Zhu and J. Wang, "Power-efficient spatial reusable channel assignment scheme in WLAN mesh networks," Mobile Networks and Applications, vol. 17, no. 1, pp. 53–63, 2012.
[7] W. Shu, "Quantum-inspired genetic algorithm based on simulated annealing for combinatorial optimization problem," International Journal of Distributed Sensor Networks, vol. 5, no. 1, pp. 64–65, 2009.
[8] N. Meghanathan, "Stability-based and energy-efficient distributed data gathering algorithms for wireless mobile sensor networks," Ad Hoc Networks, vol. 19, pp. 111–131, 2014.
[9] W. Bober and C. J. Bleakley, "BailighPulse: a low duty cycle data gathering protocol for mostly-off Wireless Sensor Networks," Computer Networks, vol. 69, pp. 51–65, 2014.
[10] D. Ebrahimi and C. Assi, "Compressive data gathering using random projection for energy efficient wireless sensor networks," Ad Hoc Networks, vol. 16, pp. 105–119, 2014.
[11] G. P. Gupta, M. Misra, and K. Garg, "Energy and trust aware mobile agent migration protocol for data aggregation in wireless sensor networks," Journal of Network and Computer Applications, vol. 41, no. 1, pp. 300–311, 2014.
[12] I. Sim, K. Choi, K. Kwon, and J. Lee, "Energy efficient cluster header selection algorithm in WSN," in Proceedings of the IEEE International Conference on Complex, Intelligent and Software Intensive Systems (CISIS ’09), pp. 584–587, March 2009.
[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] J. Ma, J. Lu, and G. Q. Zhang, "Decider: a fuzzy multi-criteria group decision support system," Knowledge-Based Systems, vol. 23, no. 1, pp. 23–31, 2010.
[15] W. Huang and L. Ding, “The shortest path problem on a fuzzy time-dependent network,” *IEEE Transactions on Communications*, vol. 60, no. 11, pp. 3376–3385, 2012.

[16] J. M. Mendel, R. I. John, and F. Liu, "Interval type-2 fuzzy logic systems made simple," *IEEE Transactions on Fuzzy Systems*, vol. 14, no. 6, pp. 808–821, 2006.

[17] I. Gupta, D. Riordan, and S. Sampalli, "Cluster-head election using fuzzy logic for wireless sensor networks," in *Proceedings of the 3rd Annual Communication Networks and Services Research Conference*, pp. 255–260, Halifax, Canada, May 2005.

[18] H. Bagci and A. Yazici, "An energy aware fuzzy approach to unequal clustering in wireless sensor networks," *Applied Soft Computing Journal*, vol. 13, no. 4, pp. 1741–1749, 2013.

[19] I. Gupta, D. Riordan, and S. Sampalli, "Cluster-head election using fuzzy logic for wireless sensor networks," in *Proceedings of the 3rd Annual Communication Networks and Services Research Conference*, pp. 255–260, Nova Scotia, Canada, May 2005.

[20] J.-M. Kim, S.-H. Park, Y.-J. Han, and T.-M. Chung, “CHEF: cluster head election mechanism using Fuzzy logic in wireless sensor networks,” in *Proceedings of the 10th International Conference on Advanced Communication Technology (ICACT ’08)*, pp. 654–659, February 2008.

[21] G. Ran, H. Zhang, and S. Gong, “Improving on LEACH protocol of wireless sensor networks using fuzzy logic,” *Journal of Information & Computational Science*, vol. 7, no. 3, pp. 767–775, 2010.

[22] N. A. Torghabeh, M. R. Akbarzadeh, and M. H. Yaghmaee, “Cluster head selection using a two-level fuzzy logic in wireless sensor networks,” in *Proceedings of the 2nd International Conference on Computer Engineering and Technology (ICCET ’10)*, pp. 2357–2361, 2010.

[23] R. Zhu, “Intelligent collaborative event query algorithm in wireless sensor networks,” *International Journal of Distributed Sensor Networks*, vol. 2012, Article ID 728521, 11 pages, 2012.

[24] R. Halke and V. A. Kulkarni, “En-LEACH routing protocol for wireless sensor network,” *International Journal of Engineering Research and Applications*, vol. 2, no. 4, pp. 2099–2102, 2012.

[25] M. Novey and T. Adali, “Complex ICA by negentropy maximization,” *IEEE Transactions on Neural Networks*, vol. 19, no. 4, pp. 596–609, 2008.

[26] Y. M. Tashtoush and M. A. Okour, “Fuzzy self-clustering for wireless sensor networks,” in *IEEE/IFIP International Conference on Embedded and Ubiquitous Computing (EUC ’08)*, vol. 1, pp. 223–229, December 2008.
Submit your manuscripts at
http://www.hindawi.com