Improved algorithm of cluster-based routing protocols for agricultural wireless multimedia sensor networks

Zhang Fu1,2*, Liu Hongmei1, Wang Jun1, Qiu Zhaomei1, Mao Pengjun1, Zhang Yakun1

(1. College of Agricultural Equipment Engineering, Henan University of Science and Technology, Luoyang 471003, China; 2. Collaborative Innovation Center of Machinery Equipment Advanced Manufacturing of Henan Province, Luoyang 471003, China)

Abstract: Low Energy Adaptive Clustering Hierarchy (LEACH) is a routing algorithm in agricultural wireless multimedia sensor networks (WMSNs) that includes two kinds of improved protocol, LEACH_D and LEACH_E. In this study, obstacles were overcome in widely used protocols. An improved algorithm was proposed to solve existing problems, such as energy source restriction, communication distance, and energy of the nodes. The optimal number of clusters was calculated by the first-order radio model of the improved algorithm to determine the percentage of the cluster heads in the network. High energy and the near sink nodes were chosen as cluster heads based on the residual energy of the nodes and the distance between the nodes to the sink node. At the same time, the $K$-means clustering analysis method was used for equally assigning the nodes to several clusters in the network. Both simulation and the verification results showed that the survival number of the proposed algorithm LEACH-ED increased by 66%. Moreover, the network load was high and network lifetime was longer. The mathematical model between the average voltage of nodes ($y$) and the running time ($x$) was concluded in the equation $y=-0.0643x+4.3694$, and the correlation coefficient was $R^2=0.9977$. The research results can provide a foundation and method for the design and simulation of the routing algorithm in agricultural WMSNs.

Keywords: wireless sensor networks, routing protocol, LEACH algorithm, improved algorithm, cluster head, $K$-means clustering

DOI: 10.3965/j.ijabe.20160904.2261


1 Introduction

Various types of monitoring object information have been detected, perceived, and acquired in real time by a variety of integrated micro sensors of wireless sensor networks (WSNs) and processed by the embedded system. Sensory information was transmitted to the user terminal by a random self-organizing wireless network. A WSN is suitable for monitoring agricultural environment compared with traditional network. As the agricultural monitoring environment has become increasingly complex, rich information of the image, audio, and video is available through media. These features are urgently needed to introduce a sensor networks based on environmental test activities, which could achieve finer and accurate information of agricultural monitoring environment[1,2]. WSNs are unable to meet the precise requirements of information for environmental monitoring. To satisfy the requirement of agricultural development, the multimedia network of service should be established to collect and process multi-data at the same time. A wireless multimedia sensor network is
Based on WSN, combined with the information of audio, video, images, and a variety of scalar data to constitute wireless network. One characteristic that WSNs do not share is that a wireless multimedia sensor network could process data for audio, video, images, and other types of media information. This advantage enables wireless multimedia sensor networks (WMSNs) to have great potential applications in agriculture detection. Energy consumption is mainly concentrated on the process of sending and receiving wireless data in the WSN. Unlike WSNs, the sensor and the sink nodes of the wireless multimedia sensor networks are used to collect and process large amounts of multimedia data. Therefore, energy consumption of multimedia data acquisition and processing is about 2.5 times than that of the wireless transceiver in a wireless multimedia sensor network. Moreover, the energy consumption of wireless multimedia sensor networks is several times or even dozens of times higher than that of WSNs. Majority research have focused on algorithm and network structure based on reducing energy consumption and extending the network lifetime of WMSNs. The lifetime of WMSNs is dependent on the lifetime of each sensor node. One of the most important problems in the whole network is achieving balance in node energy consumption for maximally extending network lifetime. Network routing protocol plays an important role in WMSNs. In designing a network routing protocol with smaller delay, higher throughput, lower energy consumption, and higher QoS (Quality of Service) are needed. Such technology could further push forward the national economy.

The LEACH algorithm is a low-power, adaptive clustering routing protocol algorithm. The basic idea is to select randomly the cluster head nodes in terms of circulation. A layered structure is applied to data transmission in the LEACH algorithm, that is, ordinary nodes send collected information to a cluster head node. The information from ordinary nodes is fused by the cluster head and then conveyed to the sink node. This method could significantly reduce energy consumption of communication, because the cluster head is the only node that has to communicate with the base station. Normal nodes transmit to the cluster head only in the clustering routing protocol. In addition, in a network with a shared channel in a TDMA (Time Division Multiple Access) mechanism, a core issue is the effective allocation of channel time to provide service guarantees for flows. Except for specific timeframe, which is assigned by cluster head, the ordinary node remains inactive. However, the cluster head has been in active condition to receive information constantly. Upon completion of each round of operation, the network routing is reinitialized, including the formation of cluster and electing of cluster head.

Although the LEACH algorithm has many advantages, several deficiencies also exist. First, while the LEACH algorithm could determine the number of cluster heads by the percentage of cluster head nodes $P$, ensuring the rationality of $P$ values is difficult when using this algorithm. This limitation stems from the set being artificial. Second, the random choice method of cluster head nodes has a negative influence as it can heighten the possibility of unbalanced energy dissipation in the network. Third, the residual energy of nodes and the distance between the sensor and the sink nodes are not considered when selecting the cluster head node. Hence, the cluster head could accelerate network death if the low-power node acts as the cluster head node, which can then lead to the failure of the entire network. Meanwhile, the range of the cluster head overlay determines the number of the cluster nodes. The location distribution of nodes is not uniform, which leads to poor performance. Therefore, the research expects that the number of cluster nodes is controlled within the reasonable range.

The LEACH-ED algorithm is presented in this study to solve the aforementioned weaknesses of the LEACH algorithm. First, the optimal number of clusters was calculated by using the first-order radio model of improved algorithm to determine the percentage of the cluster head in the network. Second, the K-means clustering method, instead of the randomizing method, was used to select cluster heads. The proposed method proved to be an effective way to solve such problems. The residual energy of nodes and the distance between the normal nodes and the sink nodes could be combined.
2 Mathematical models

2.1 Calculating the optimal number of cluster heads

The number of cluster heads has a significant influence on the entire network’s lifetime\cite{10-13}. In this study, the percentage P of the cluster head\cite{14-17} was calculated in control of the optimal number of the cluster head by using the energy adaptive cluster head selection algorithm. Assume 𝑁 nodes and 𝑄 cluster head nodes in 𝑀*𝑀 monitor environment. Each cluster contains 𝑁/𝑄 nodes, including a cluster head and (𝑁/𝑄)-1 ordinary node.

According to the first-order radio model\cite{19,20}, the energy consumption of sending 1 bit data could be worked out by using the relation

\[ E_{\text{send}} = E_{\text{elec}} + E_{\text{amp}} \cdot d^\beta \]  \hspace{1cm} (1)

Where, \( E_{\text{elec}} \) is the energy dissipated for the radio dissipates to run the transmitter or receiver circuitry, and \( E_{\text{amp}} \) is the energy dissipated of transmit amplifier.

We could obtain the corresponding energy consumption of receiving 1 bit data by using the equation

\[ E_{\text{receive}} = E_{\text{elec}} \]  \hspace{1cm} (2)

Therefore, the communication energy consumption formula between cluster head nodes and sink node is defined by

\[ E_{\text{CH}} = x \cdot \left( \frac{N}{K} - 1 \right) \cdot E_{\text{elec}} + x \cdot \frac{N}{K} \cdot E_{\text{DA}} + x \cdot E_{\text{amp}} \cdot d_{\text{toCH}}^\beta \]  \hspace{1cm} (3)

where, \( x \) is the number of sending bits. The multipath fading channel mode was used for communicating with sink node and cluster nodes, and the value of variable \( \beta \) was set to 4. In addition, \( E_{\text{DA}} \) is the energy consumption of the cluster nodes for data fusion, and \( d_{\text{toCH}} \) is the distance between the cluster and the sink nodes.

Usually, the ordinary nodes and the cluster head have short distances in the cluster. The free space channel model was used for initiating communication between the two kind of the nodes, and \( \beta \) was set to 2. The process of transmitting \( x \) bit data package needs to spend energy, and the equation is given by

\[ E_{\text{FM}} = x \cdot E_{\text{elec}} + x \cdot E_{\text{amp}} \cdot d_{\text{toBS}}^2 \]  \hspace{1cm} (4)

where, \( d_{\text{toBS}} \) is the distance between the ordinary node and the cluster node.

Let us suppose that both the cluster heads and the sensor nodes are equally distributed in the entire monitor area, and the whole monitor area is contained in a circular area. The coverage radius of the cluster node and the distribution density of the nodes were calculated based on the assumption that the monitor area of each cluster was \( M^2/K \). The coverage radius of cluster head node was \( R = M/\sqrt{\pi K} \), and the distribution density of the nodes was \( \rho(X,Y) = K/M^2 \). Hence, the sum of squared of the distance between the ordinary nodes and the cluster heads in the cluster is defined as

\[ d_{\text{toCH}}^2 = \iint (X^2 + Y^2) \cdot \rho(X,Y) dX dY \]

\[ = \frac{\rho \cdot M^4}{2\pi K^2} = M^2 \]  \hspace{1cm} (5)

A cluster needs to consume energy \( E_{\text{CW}} \) in a round, and the total energy consumption for the 𝑄 clusters in the entire monitoring environment is \( E_{\text{total}} \), which is expressed as

\[ E_{\text{total}} = Q \cdot E_{\text{CW}} = Q \cdot E_{\text{CH}} + K \cdot \left( \frac{N}{K} \right) \cdot E_{\text{FM}} \]

\[ = x \cdot \left[ 2N \cdot E_{\text{elec}} - K \cdot E_{\text{elec}} + N \cdot E_{\text{DA}} + K \cdot E_{\text{amp}} \cdot d_{\text{toCH}}^4 + N \cdot E_{\text{amp}} \cdot \frac{M^2}{2\pi K} \right] \]  \hspace{1cm} (6)

A simplified mathematical model of the optimal cluster head was established according to the optimal target with the minimal energy consumption in the entire network. In this case, the equation is given by

\[ K = \frac{M}{d_{\text{toCH}}^2} \cdot \sqrt{\frac{N \cdot E_{\text{amp}}}{2\pi e_{\text{amp}}}} \]  \hspace{1cm} (7)

Then, \( P=K/N \), where \( N \) is the number of total nodes in the network.

The main target of the improved routing algorithm was to maintain the energy usage of the sensor nodes efficiently. A model for the number of the cluster head was set up, and a simulation was carried out under the setting condition in the monitoring environment. The monitor area was 100×100, the number of nodes was 100, and the base station location was (50, 50). Each sensor
node was assumed to have an initial energy of 0.5 J. The experiment was performed (shown in Table 1). Table 1 lists values associated with each of the named variables. Finally, the optimal theoretical number of cluster head was $2 \leq K \leq 4$. The simulation of the actual measurement data and the result of experiments were analyzed. The different numbers of the cluster heads and the average residual energy diagram of each round were also finalized, as shown in Figure 1.

Considering the number of cluster heads as 4, the larger average residual energy of the node, the less the simulation process consumed total energy (Figure 1). After that, the number of cluster heads was set to 4.

### Table 1  Simulation Parameters

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monitor area/m</td>
<td>100×100</td>
</tr>
<tr>
<td>Number of nodes</td>
<td>100</td>
</tr>
<tr>
<td>Base station location/m</td>
<td>(50, 50)</td>
</tr>
<tr>
<td>Initial energy/J</td>
<td>0.5</td>
</tr>
<tr>
<td>Packet length/bit</td>
<td>4000</td>
</tr>
<tr>
<td>$E_{elec}$/J</td>
<td>$5 \times 10^{-6}$</td>
</tr>
<tr>
<td>$E_{sta}$/J</td>
<td>$5 \times 10^{-9}$</td>
</tr>
<tr>
<td>$\varepsilon_{fs}$/J</td>
<td>$10^{-11}$</td>
</tr>
<tr>
<td>$\varepsilon_{amp}$/J</td>
<td>$1.3 \times 10^{-15}$</td>
</tr>
</tbody>
</table>

Figure 1 Relationship curve between the cluster heads and the average remaining energy

2.2 Cluster division

A cluster is typically divided using a random fashion, which leads to a largely uneven division. The balanced or unbalanced node location distribution can directly affect the performance and lifetime of the network.

In this study, the K-means clustering algorithm was applied to the LEACH algorithm to divide the clusters and make them as balanced as possible. This was also done to reduce energy consumption and communication delay of the nodes [18-23].

The distance between two points is calculated by using the equation

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$  \hspace{1cm} (8)

where, $(x_1, y_1)$ and $(x_2, y_2)$ represent the coordinates of the two points, respectively.

Here, we present details of the procedure. To divide the whole network into multiple clusters, the distance between one node and all nodes was calculated by using the K-means clustering algorithm in the network based on the entire network coordinate values. The LEACH algorithm followed this specific process. The effects of the cluster when the sink nodes are (0, 0) and (50, 50) are shown in Figures 2 and 3, respectively.

Figure 2  Division cluster effect of the entire network distance with the sink coordination (0, 0)

Figure 3  Division cluster effect of the entire network distance with sink coordination (0, 0)

Regardless of the distance between the nodes and the cluster head used by the K-means clustering algorithm within the monitored area of the sink node, the algorithm ensures uniform distribution of the cluster in the entire
monitoring area and the number of members for cluster stability.

The new algorithm has significant advantages in terms of communication time delay, because its nodes could communicate with the cluster head without multi-hop in the data transmission stage, as shown in Figures 2 and 3, respectively.

2.3 Selecting the cluster head

The nodes at the edge of monitoring area or the low energy nodes were selected as the cluster head node. However, this condition could lead to the failure of the cluster if the distance between nodes and node energy consumption was not considered in the LEACH algorithm. This research presents the LEACH-ED algorithm, which introduces node energy and distance to the cluster head selection mechanisms. The improved threshold is expressed as

\[
T(n) = \begin{cases}
    P \times \left( \frac{E(i)r}{E_{int}} \times (1 - D(i)) \right) & n \in G \\
    0 & n \notin G
\end{cases}
\]

\[
D(i) = 0.1 \times D_1(i) + 0.9 \times D_2(i)
\]

where, \(E(i)r\) is the residual energy of the node \(i\) to work after \(r\) rounds; \(E_{int}\) is the initial energy of the nodes; \(D_1(i)\) is the sum of distance from node \(i\) to another node in the cluster; \(D_2(i)\) is the distance from node \(i\) to the sink node, and \(D(i)\) is the distance from node \(i\) to other nodes. The value of distance from each node to the sink node must be normalized when the distance fluctuation is large. The normalized result is recorded in \(D_1(i), D_2(i), D(i)\).

3 Simulation method

This study first presents the conception and the process of clustering algorithms. Then, the improved algorithm is designed and presented to solve the issue of optimal cluster heads. The mathematical models of the LEACH-ED, LEACH-C, LEACH-D, and LEACH algorithms are established by analogue simulation computation with Matlab software. The simulation method randomly generates 100 sensor nodes in 100×100 areas of coverage and sets sink node coordinates (50, 50). The four kinds of threshold \(T(n)\) calculation formula are listed in Table 2.

### Table 2  Four kinds of threshold \(T(n)\)

<table>
<thead>
<tr>
<th>Algorithm name</th>
<th>Formula of threshold (T(n))</th>
</tr>
</thead>
</table>
| LEACH          | \[
T(n) = \begin{cases}
    P \times \left( \frac{E(i)}{E_{int}} \times (1 - D(i)) \right) & n \in G \\
    0 & n \notin G
\end{cases}
\] |
| LEACH-E        | \[
T(n) = \begin{cases}
    P \times \left( \frac{E(i)r}{E_{int}} \right) & n \in G \\
    0 & n \notin G
\end{cases}
\] |
| LEACH-D        | \[
T(n) = \begin{cases}
    P \times \left( \frac{E(i)r}{E_{int}} \times (1 - D(i)) \right) & n \in G \\
    0 & n \notin G
\end{cases}
\] |
| LEACH-ED       | \[
T(n) = \begin{cases}
    P \times \left( \frac{E(i)r}{E_{int}} \times (1 - D(i)) \right) & n \in G \\
    0 & n \notin G
\end{cases}
\] |

4 Analysis of simulation results

4.1 Comparison analysis on the number of nodes

A WMSN has a fault tolerance. Hence, the survival time of the nodes is among the most important factors in ensuring the quality of the entire network. In this study, the network began to enter the death state with the death of the first node. The network traffic become increasingly poor and 20% of the nodes eventually died. The mortality rate of the nodes went up by 50%, which decreased the monitor area of network. Consequently, the network died when 80% of the nodes died. Energy saving is the focus of many studies in this area because of the limited energy of a single node. Here, the number of surviving nodes was calculated by using the LEACH-ED, LEACH-D, LEACH-E, and LEACH algorithms in the network. The calculation results are shown in Table 3.

### Table 3  Contrast table of four LEACH algorithms between the protocols and the death time of the nodes

<table>
<thead>
<tr>
<th>Algorithm name</th>
<th>Death node number</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First node</td>
</tr>
<tr>
<td>LEACH</td>
<td>761</td>
</tr>
<tr>
<td>LEACH-E</td>
<td>980</td>
</tr>
<tr>
<td>LEACH-D</td>
<td>828</td>
</tr>
<tr>
<td>LEACH-ED</td>
<td>1065</td>
</tr>
</tbody>
</table>

The issues and their associated protocols were discussed and shown in Figure 4. As can be seen, the number of dead nodes was directly proportional to running time. The comparison research on run time of network was performed by using two algorithms under specific horizontal conditions. Experimental results
showed that, between the two data sets and clustering protocol (LEACH and LEACH-ED), the LEACH-ED algorithm, in the best case, increased sensor network lifetime by 140%, 150%, 160%, and 340% in the four states, respectively. Moreover, the reliability information of the transmission in LEACH-ED was higher than that in LEACH. Similarly, the relations among LEACH, LEACH-D, and LEACH-E were correctly processed. The test experiment data were analyzed and summarized based on the four kinds of algorithms, including the LEACH, LEACH-D, LEACH-E, and LEACH-ED algorithms. The number of network operating rounds with the horizontal and the vertical survival node number for the curve of the surviving node number are shown in Figure 4.

![Figure 4 Curves of the surviving node numbers](image)

### 4.2 Comparison analysis on residual energy

Residual energy reflects the energy consumption of a WMSN. The greater the amount of residual energy, the less energy consumption for data sending, receiving, and processing will be, and the longer the network lifetime. The calculation results for residual energy, which describe the relevant changes in the curves of the network, are indicated in Table 4 and Figure 5.

![Figure 5 Curves of the residual energies](image)

The number of rounds for network operation was used as the x-axis, and the residual energy of nodes was used as the y-axis. In comparing the curves for the four algorithm results, we found that the curve of the improved LEACH-ED algorithm was significantly higher than those of the three other algorithms when the linear range of horizontal was 500-4000. The results also showed that the network residual energy of the LEACH-ED algorithm was obviously higher than those of the three other algorithms under the specified subsection knots of horizontal conditions. The LEACH-ED algorithm increased the residual energy of the nodes by 107%, 193%, and 934%, respectively. The remaining energy was roughly equal to four algorithms outside 500-4000. Hence, the LEACH-ED algorithm was effective in terms of extending the network lifetime.

### 5 Experimental verification

This study proposes an improved algorithm, the LEACH-ED algorithm, which is based on the LEACH-D, LEACH-E, and LEACH algorithms. Simulation experiments were conducted to validate the superiority of the algorithm performance. Further validation of the simulation results is necessary, and other scholars have proven that the performances of both the LEACH-D and LEACH-E algorithm are better than that of LEACH[14,15]. As duplication is no longer needed, a comparison of the LEACH-ED and LEACH algorithm performances should be done. However, arranging hundreds of sensor nodes in limited experimental conditions is difficult. The best solution is to select the representative positions with which...
to arrange the node, just like described below.

The verification test is also required to evaluate the reliability of the network. Test conditions were set as follows: transmission power of nodes was 9.5 dBm and the aerial gain of the antenna was 45.69 dBZ with the height of 1.5 meters above ground. Then, the proving ground was established, which contained A, B, C, D, E five sensor nodes and a sink node F. Figure 6 shows information about the coordinate position of each node. Two algorithms, LEACH and LEACH-ED, were verified.

Scheme 1: A 3.6 V/8100 mAh MH/Ni battery provided power for the sensor nodes. We ensured that the node battery would still be charged before the experiment. The sink node generates a random number from 1 to 5 as the cluster head node, which represents A, B, C, D, and E nodes, respectively.

Scheme 2: Under the same experimental conditions, the distance between the sensor and the sink nodes was calculated according to the coordinates of each node in the second scheme. $T(n)$ could be calculated using the formula

$$T(n) = \frac{E(i)r}{E\text{int}} \times (1 - D(i))$$

(11)

The largest $T(n)$ of the five nodes was selected as the cluster head node.

Schemes 1 and 2 are based on the theories that support the LEACH and LEACH-ED algorithms, respectively. As shown in Figure 7, Scheme 2 was obviously better than 1 in terms of the survival number of the nodes. The survival number of the nodes increased by 66% in the second scheme.

The node average voltage trended down, as shown in Figure 8. However, a significant difference between the two schemes was observed based on the voltage curves. The average voltage curve of the second scheme gently declined compared with the first scheme sharp decline curve.

The relationship between the average voltage of nodes ($y$) and the running time ($x$) could be well described by the following linear equation:

$$y = -0.0643x + 4.3694$$

The correlation coefficient was
The energy consumption in the second scheme was only about half of that in first scheme. Therefore, the LEACH-ED algorithm significantly reduced energy consumption and effectively extended the lifetime of the network.

6 Conclusions

An improved LEACH-ED algorithm was proposed to solve the shortage of LEACH-based algorithms. The characteristics and evaluation standard of cluster-based routing protocols for agricultural wireless multimedia sensor networks and the LEACH algorithm were analyzed.

1) We assumed that \( N \) sensors and \( K \) cluster heads nodes were uniformly dispersed in the \( M\times M \) square region. The first-order radio model was used to calculate the optimal number of the cluster head and the percentage of network cluster head according to the monitor environment. Simulation results showed that the calculation approached the optimal requirements.

2) Cluster heads always transfer data to a base station, which is often far from sensor nodes. Hence, the \( K \)-means clustering analysis method was used in the LEACH-ED algorithm. By using the proposed clustering method, the nodes in the network can be divided into several clusters according to the distance between the nodes and the optimal number of cluster heads can be determined.

3) Residual energy and the distance between nodes were preferentially considered when choosing the cluster heads. Simulation results showed that the sensor network lifetime and the residual energy of nodes increased by 140% and 107%, respectively.

4) Both simulation and verification results show that the improved algorithm LEACH-ED is superior to other algorithms in energy-balancing and the network lifetime, which can improve network performance including network lifetime and processing speed in WMSNs for future environmental monitoring and management in agriculture.

Acknowledgements

This work was supported by Project on the Integration of Industry, Education and Research of Henan Province (Grant No. 142107000055, 162107000026), Scientific and Technological Project of Henan Province (Grant No. 152102210190, 162102210202), Natural Science Foundation of Henan Educational Committee (Grant No. 14B416004, 14A416002 and 13A416264), Key Project of Henan Tobacco Company (HYJK201316) and Innovation Ability Foundation of Natural Science (Grant No. 2013ZCX002) of Henan University of Science and Technology.

[References]


