

Non-uniform clustering routing protocol of wheat farmland based on effective energy consumption

Yisheng Miao^{1,2}, Chunjiang Zhao^{2,3*}, Huarui Wu^{2,3}

(1. College of Computer Science, Beijing University of Technology, Beijing 100124, China;

2. National Engineering Research Center for Information Technology in Agriculture, Beijing 100097, China;

3. Beijing Research Center for Information Technology in Agriculture, Beijing Academy of Agriculture and Forestry Sciences, Beijing, 100097, China)

Abstract: Wireless sensor network (WSN) can achieve real-time data collection and transmission of environment, soil, meteorology, crop physiology and other information in agriculture. The data provided by WSN could be used for decision making and management, which is very important in precision agriculture. Wheat farmland wireless sensor network has the characteristics of wide coverage area, long planting period, inconvenient energy supply, and serious impact of crop environment on wireless signal transmission. Routing protocol is an important method to achieve long-term WSN monitoring by selecting an appropriate path with low energy consumption for data transmission. According to the phenomenon of uneven environment and channel parameters caused by intensive crop growth in farmland, a non-uniform clustering routing protocol based on effective energy consumption (UCEEC) was proposed in this work. The method combined with the characteristics of multi-path fading of farmland environment signals. The idea of image segmentation was introduced. Nodes with high similarity were divided into a cluster area by the dissimilarity between nodes in order to improve the intracluster communication performance. Meanwhile, a multi-hop path selection method between cluster-heads based on the estimation of two-hop effective energy consumption is designed. The energy consumption cost factor is calculated by the effective energy consumption and the average energy consumption within the cluster to achieve the minimum and balance of the overall energy consumption of the network. Simulation results show that, compared with the existing Maximum Residual Energy Based Routing (MREBR) protocol, minimum Energy Consumption Based Routing (MEC) routing protocols, UCEEC improves the energy balance effect between nodes, prolongs the network life cycle, and realizes efficient energy utilization of wireless sensor network data collection in the complex environment of wheat field.

Keywords: farmland, wireless sensor networks, multi-path fading channel, routing protocol, energy optimization

DOI: 10.25165/j.ijabe.20211403.6540

Citation: Miao Y S, Zhao C J, Wu H R. Non-uniform clustering routing protocol of wheat farmland based on effective energy consumption. *Int J Agric & Biol Eng*, 2021; 14(3): 163–170.

1 Introduction

Precision agriculture can effectively increase crop yield and quality by adjusting the environment, water and fertilizer conditions of crop growth. It can also reduce the use of agricultural resources, reduce production costs, and improve ecological environment. Internet of Things (IoT) technology can achieve real-time data collection and transmission of farmland environment, soil, meteorology, crop physiology, and other information, and provide fundamental data sources for intelligent decision-making and production management. Monitoring data analysis can effectively help to improve the yield and quality of agricultural products, which is an important technical method to achieve precision agriculture. Wireless sensor network (WSN) is an important IoT application. Because of its wireless,

self-organization, low cost, low power consumption, and other characteristics, it has been widely used in industry, medical, agriculture, and other fields. Large-scale farmland monitoring network has many characteristics, such as wide coverage area, long planting period, inconvenient energy supply, complex and changeable environment, and etc. How to achieve stable, reliable, and efficient data transmission and collection is a difficulty in research and application^[1,2]. As an important part of network data collection application, routing protocol has a great impact on network performance, and it is one of the hotspots of WSN network data collection research in recent years^[3,4].

Stable network communication is the foundation of wireless sensor network monitoring application in farmland. In practical application, the working condition of wireless communication is harsh. Wireless signal would not only disperse and lose with the increase of propagation distance, but also the luxuriant branches, leaves and fruits of crops have significant shielding effect on the radio signal. The influence of crop growth on wireless signal propagation must be fully considered^[5]. Therefore, it is necessary to study the energy-efficient routing method under the condition of crop growth.

The free space fading model and the shadow fading model are commonly used in communication field^[6]. The free space attenuation model is an ideal model, which is only suitable for completely open area. The shadow fading model only considers the influence of terrain^[6]. These two models do not consider the

Receive date: 2021-02-22 **Accepted date:** 2021-03-30

Biographies: **Yisheng Miao**, PhD candidate, Associate Researcher, research interest: IoT and artificial intelligence for agricultural application, Email: yishengmiao@126.com; **Huarui Wu**, Researcher, research interest: big data analysis and intelligent decision for agricultural application, Email: wuhr@nercita.org.cn.

***Corresponding Author:** **Chunjiang Zhao**, Researcher, PhD Advisor and Academician of Chinese Academy of Engineering, research interest: intelligent agriculture, machine learning and agricultural information technology. Beijing Nongke Mason, 11#Shuguang Huayuan Middle Road, Haidian District, Beijing 100097, China. Email: zhaocj@nercita.org.cn.

effect of dense crop occlusion on signal propagation, and are not suitable for describing the small-scale fading phenomenon caused by the reflection, scattering and absorption of wireless signals by crops in farmland environment. Many studies have been done on wireless channel modeling in agricultural environment. Vougioukas et al.^[7] carried out a wireless signal attenuation experiment of orchard Internet of things. Wang et al.^[8] studied the production control technology based on wireless sensor network for greenhouse microclimate environment. Guo et al.^[9] studied the relationship between antenna height, environmental parameters, and wireless signal strength attenuation through an apple orchard experiment. These studies are dominated by large-scale fading, without considering the small-scale fading effect of multipath signal caused by crop occlusion. Miao et al.^[5] studied the small-scale fading phenomenon of signals in wheat field under non line of sight (NLOS) condition. A multi-scale fading channel model based on statistical distribution was proposed, which provides a farmland channel model in this study.

In terms of WSN routing algorithm, the clustering routing algorithm has the advantages of a simple structure and efficient topology management, which is more suitable for resource constrained farmland wireless sensor networks. In order to improve the energy efficiency and balance of network nodes and prolong the network life cycle, a series of improved routing methods have been studied. Minimum energy consumption (MEC)^[10] routing designed the protocol based on minimum energy consumption of data transmission, without considering the residual energy of nodes which are outside the path of minimum energy consumption in the network, so the nodes on the minimum energy path will die early due to heavier forwarding tasks^[11,12]. Self-organizing routing (SOR) is an energy aware protocol. A payment strategy is introduced to motivate nodes to undertake data transfer tasks, and sensor nodes make routing decisions based on local information. The protocol achieves the combination of effectiveness and equalization of energy consumption. In the clustering process, EECS (Energy Efficient Clustering Scheme)^[14] transmitted the data of the cluster-head to the sink node through single-hop communication. Ordinary nodes chose the appropriate cluster-head according to the distance between itself and the cluster-head and the distance between the cluster-head and the sink node, and a non-uniform cluster is constructed. In this way, the clusters that farther away from the sink node have smaller size and it balanced energy consumption within the clusters. However, this solution can only balance the energy consumption among cluster-heads and it cannot achieve overall energy balance^[15,16]. UCS (Uneven Clustering Scheme)^[17] uses the idea of non-uniform clustering to balance the energy consumption of cluster-head, and the multi-hop transmission between clusters was adopted. The energy consumption of the cluster head is mainly composed of two parts, namely the intra-cluster communication and the inter-cluster transmission. The energy consumption of cluster communication is proportional to the number of member nodes. The energy consumption among clusters mainly caused by the amount of data transmitted. The algorithm adjusted the cluster size according to the load that the cluster head forwarding data, so that the energy consumption of the cluster heads are close to each other and the energy consumption of the entire network is balanced^[18-20].

The above clustering algorithms regard the network links as reliable connected links in the routing process, and estimate the transmission energy consumption between nodes according to the large-scale fading model. Clustering is carried out without

considering the multi-scale fading of farmland environment. Chen Yang et al.^[21] reduced the impact of crop growth on network connectivity through backup routing, and realized reliable data transmission in a greenhouse. Pandiyaraju et al.^[22] combined network energy efficiency with agricultural irrigation, and selected data transmission path in precision agriculture environment by introducing a fuzzy reasoning system, which improved network lifetime. In addition, signal strength, quality of service parameters, network power efficiency are also considered in agricultural WSN data transmission routing. The related research shows that the communication qualities of nodes in different crop growing areas are different due to the different planting and growing density^[23-25]. By analyzing the strengths of the signals of sensor nodes, nodes are clustered according to the idea of image segmentation, and nodes that in the same connected area and ventilation are classified as a cluster. Based on the clustering, a novel routing protocol based on environmental factors and two greedy algorithms are designed^[26-30].

In summary, when the above routing protocols are applied to the complex environment of farmland, most of them do not consider the environmental characteristics and network energy consumption at the same time. Therefore, this work proposes a novel non-uniform clustering routing protocol. The network nodes are divided into clusters according to the characteristics of the environment channel. The network energy performance is improved via multi-hop path selection between cluster-heads. The overall network energy consumption is reduced and balanced. The WSN life cycle in the complex environment of wheat farmland is extended.

2 System model and assumptions

2.1 Network model

In this study, a farmland wireless sensor network scenario is proposed^[31-33]. It is assumed that N sensor nodes are randomly deployed in a two-dimensional rectangular region of $X \times Y$, and the wireless sensor network is assumed to have the following properties:

- 1) The network is a static network with higher density, which means the deployment of sensor nodes remains unchanged, the node density is enough to ensure network connectivity and coverage of the monitoring area.
- 2) Sink node is fixed and unique with unlimited energy supply, and its wireless transmitting power is controllable.
- 3) The sensor nodes are isomorphic and non-rechargeable. The initial energy is the same as E_0 .
- 4) The energy consumed by each sensor node is not equal, which makes the network energy heterogeneous.
- 5) Nodes have the ability of self-energy perception.
- 6) The nodes have the ability of data fusion.

2.2 Energy model

The network wireless communication power model was set as described in reference [5]. All nodes had two types of transmitting amplifier power, higher one for long distance mode and lower one for near distance mode. The node transmission power E_{tx} is calculated as:

$$E_{tx} = \begin{cases} lE_{elec} + l\epsilon_f d^2, & d < d_{crossover} \\ lE_{elec} + l\epsilon_m d^{4+x}, & d \geq d_{crossover} \end{cases} \quad (1)$$

where, l refers to the number of transmitted bits; E_{elec} refers to the 1 bit data' transmission power of the transmit-receive circuits, dBm; d refers to the distance between the transmitter and receiver, m;

$d_{crossover} = \sqrt{\varepsilon_f / \varepsilon_m}$ refers to the distance threshold of the model; ε_f and ε_m refer to the energy coefficients for the power amplification in the two conditions; x refers to the fast fading component in the multi-path propagation.

Nodes required energy for receiving l bit data, the node receiving power E_{rx} could be calculated as below,

$$E_{rx} = lE_{elec} \quad (2)$$

By Equation (1), in the near distance communication, the transmission attenuation energy consumption is proportional to the square of the distance, while in the long distance communication, it is proportional to the 4th power of distance. Because of the frequent communication and large amount of data, the one hop communication radius is defined as $d_{crossover}$ according to Equation (1).

3 Non-uniform clustering routing protocol based on effective energy consumption

3.1 Node correlation based clustering

A hybrid routing composed of local clustering and multi-hop routing between cluster-heads were adopted. Non-uniform clustering based on signal channel similarity in local area could achieve stable intra-cluster connection, which would improve network energy performance. A multi-hop mode can reduce the extra energy consumption of cluster-heads far away from the sink node, and realize the energy balance of cluster-heads in different regions.

Among current methods, the clustering algorithm of node association degree clusters nodes based on distance. Then the cluster heads are selected according to the residual energy of nodes^[34-36]. The advantage of dividing the nodes with high correlation into the same cluster is that the energy consumption of intra cluster transmission is lower, and the data fusion can achieve better results at the cluster-head. Under the condition of LOS (Line of Sight), it can be considered that the node correlation is negatively correlated with the distance between nodes. However, the dense growth of crops in farmland results in the occlusion barrier, which forms the non-uniform channel conditions. It can also be considered that this kind of occlusion will also cause the non-uniformity of environmental parameters. That is, because of the occlusion of crops, the correlation and distance between nodes do not present a monotonic function^[34]. Therefore, this paper proposes a non-uniform clustering routing protocol based on effective energy consumption (UCEEC). Using the idea of image segmentation^[35], a similarity region segmentation method is used in UCEEC's clustering procedure. The correlation between nodes is represented by the actual transmission distance.

In order to realize the clustering of node association degree based on a channel model, firstly, the network is regarded as an undirected graph, WSN nodes are nodes in the undirected graph, and the links between adjacent nodes are edges of the undirected graph. According to the weight of the edge, the dissimilarity between the two nodes is measured, and then whether the two nodes should belong to a region is determined. Finally, the nodes belonging to a region are clustered. Specifically, the signal fading between two nodes is taken as the weight of edges in the undirected graph, and each node is an independent region in the initial state.

The internal region spacing $\text{Int}(C)$ and inter-region dissimilarity $\text{Dif}(C_m, C_n)$ are defined as follows.

Definition 1: Internal region spacing

$$\text{Int}(C) = \max_{e \in \text{MST}(C, E)} (w(e)) \quad (3)$$

where, C is the set of nodes in the region; e is the set of edges of nodes in the region, and the interval in the region is the weight value of the edge with the largest weight in MST (minimum spanning tree) of nodes in the region.

Definition 2: Inter region dissimilarity

$$\text{Dif}(C_m, C_n) = \min_{v_i \in C_m, v_j \in C_n, (v_i, v_j) \in E} (v_i, v_j) \quad (4)$$

where, C_m and C_n represent two adjacent regions; v_i and v_j are nodes in C_m and C_n , respectively. That is to say, among all pairs of points which belong to two regions and have edge connection, the pair with the least weight is found. If the points in two regions are not connected by edges, then the distance is defined as positive infinity.

When the space between regions is smaller than that within regions, the adjacent regions can be merged. The functional form of the assertion is given in reference [35].

$$D(C_1, C_2) = \begin{cases} 1 & \text{if } (\text{Dif}(C_1, C_2) < \text{MInt}(C_1, C_2)) \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

where, $\text{MInt}(C_1, C_2) = \min(\text{Int}(C_1) + \gamma(C_1), \text{Int}(C_2) + \gamma(C_2))$, γ is a threshold function, which is used to control the extent that the distance between regions is greater than that within regions.

In order to achieve better clustering and energy consumption performance, the region similar to the circle should be made when merging regions, and avoided the strip-shaped region distribution. Therefore, the maximum distance between nodes in the merged region is limited.

$$\frac{|C|}{L_{\max}(C)} < \sigma \quad (6)$$

where, $|C|$ is the number of nodes in area C ; $L_{\max}(C)$ is the maximum distance between nodes in area C .

If $D(C_1, C_2) = 1$ and satisfies Equation (6), region merging is performed. When the specific algorithm is executed, for a node i in the network, the neighbor nodes that are not in the same region are judged in order of high to low dissimilarity between regions. After each region merging, the node set of non-common region neighbors and the corresponding domain values would be updated. After the cycle judgment, the node with the highest residual energy is elected as the cluster-head.

The pseudo code of this algorithm is as follows:

Algorithm: Node correlation based clustering algorithm

Input: The number of sensors N , channel model in wheat farmland

Output: The cluster partition of all nodes

1. Network Initialization.

2. Randomly generate initial position of all nodes. $v_i = \text{random}(x, y)$

Set the residual energy of all nodes to E_0 .

3. Set the network as an undirected graph, every node belongs to a different region C_i , where $i=1, 2, \dots, R$

Remark the state as $S(0)$.

4. Calculate all the edge weight $w(e)$ in $S(0)$. // Using MST algorithm.

Calculate the internal region spacing $\text{Int}(C)$ and inter region dissimilarity $\text{Dif}(C_m, C_n)$ in $S(0)$

$\text{Int}(C) = 0$ (at the initial state $S(0)$)

$\text{Dif}(C_m, C_n) = \min_{v_i \in C_m, v_j \in C_n, (v_i, v_j) \in E} (v_i, v_j)$

5. while all $C_i \subset S(n)$ // n is the round number of algorithm iteration

6. if $\text{Dif}(C_m, C_n) < \text{MInt}(C)$

7. Nodes in C_m and C_n merge to form a new cluster C_k

8. else if

9. Keep the C_m and C_n unchanged
10. end if
11. if all $C_i \subset S(n)$ are not changed
12. break
13. end if
14. All the regions C_i form the new state $S(n \leftarrow n+1)$
15. Calculate the internal region spacing $Int(C)$ and inter region dissimilarity $Dif(C_m, C_n)$ in new $S(n)$

$$Int(C_k) = \max_{e \in MST(C_k, E)} (w(e))$$

$$Dif(C_m, C_n) = \min_{v_i \in C_m, v_j \in C_n, (v_i, v_j) \in E} (v_i, v_j)$$

16. end while

3.2 Multi-hop between cluster heads based on effective energy consumption

In UCEEC, the multi-hop mode can reduce the extra energy consumption of cluster-heads far away from the base station (sink node), and realize the energy balance of cluster heads in different regions. The hybrid routing method is shown in Figure 1.

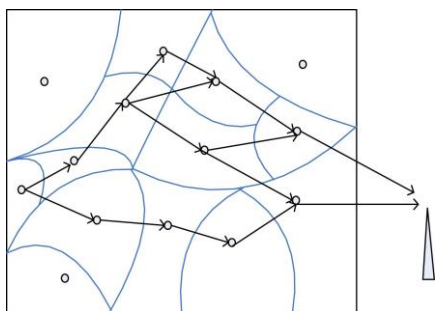


Figure 1 Multi-hop between cluster-heads based on effective energy consumption

Due to the non-uniform channel characteristics in farmland environment, the path loss between cluster-heads in different regions is significantly different. It also means that the transmission energy consumption between cluster-heads is different. The distribution of transmission energy consumption is shown in Figure 2. The distance between cluster-heads and sink node is also indicated in Figure 2, where d represents the unit distance, and e represents the unit energy consumption.

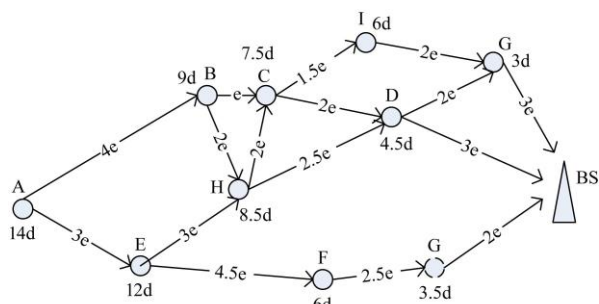


Figure 2 Distribution of transmission energy consumption

It is difficult for nodes to get the global path loss information. Therefore, this study proposes an effective displacement and energy cost method to realize multi-hop path selection based on local information. In each path selection, the path effective displacement under the unit energy cost is larger, that is, the data are closer to the target node.

As shown in Figure 3, the path loss of cluster-head A to C via B is $5e$, where e is the unit path loss. As an example, the effective displacement of path "A-C" is calculated as follows. In order to simplify the derivation, the coordinates of the target node is set as $(0, 0)$.

$$|CS| = \sqrt{x_c^2 + y_c^2} \tag{7}$$

$$\tan \angle ASA' = \frac{|AS|}{|A'S|} = \frac{x_a}{y_a} \tag{8}$$

$$|AS| = \sqrt{x_a^2 + y_a^2} \tag{9}$$

$$\tan \angle CSC' = \frac{|CS|}{|C'S|} = \frac{x_c}{y_c} \tag{10}$$

From Equations (7)-(10), the effective displacement $|A'C'|$ can be obtained as

$$\begin{aligned} |A'C'| &= |A'S| - |C'S| = \frac{|AS| \cdot y_a}{x_a} - \frac{|CS| \cdot y_c}{x_c} \\ &= \frac{\sqrt{x_a^2 + y_a^2} \cdot y_a}{x_a} - \frac{\sqrt{x_c^2 + y_c^2} \cdot y_c}{x_c} \end{aligned} \tag{11}$$

According to the path information within two hops, cluster-head A could determine the effective displacement of the possible path, calculates the corresponding energy consumption cost, and updates the neighbor cluster head information table recorded by the cluster head.

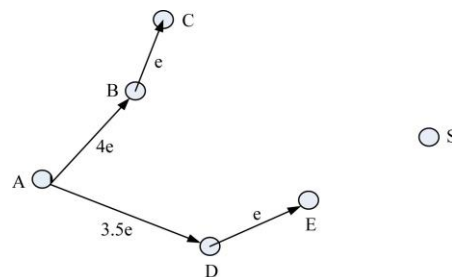


Figure 3 Calculation of effective displacement of path

In order to increase the utilization of energy consumption, the neighboring cluster head information table of the cluster head is updated according to Equation (12).

$$E_{i,j,k} = \frac{e_{i,j} + e_{j,k}}{d_{i,BS} - d_{k,BS}} \tag{12}$$

where, $E_{i,j,k}$ represents the energy consumption of node i transferring data to node k via node j ; $e_{i,j}$ represents the energy consumed that node i transferring data to node j ; $d_{i,BS}$ represents the distance from node i to sink node BS, which is calculated by the following equation:

$$d_{i,BS} = \sqrt{(x_i - x_{BS})^2 + (y_i - y_{BS})^2} \tag{13}$$

The node with minimum two-hop estimation energy consumption is selected as the next hop node.

$$E_{i,j,k} = \min(E_{i,p,q}), p \in FN(i), q \in FN(p) \tag{14}$$

where, $FN(i)$ refers to the forwarding node set of node i .

The single-hop greedy algorithm based on energy consumption is used to solve the route forwarding path, the result is $A \rightarrow E \rightarrow H \rightarrow C \rightarrow D \rightarrow BS$, whose energy consumption is $13e$. In this algorithm that is based on the energy consumption per unit distance within two hops, for each time of transferring data, the forward nodes should be found first, that is, the nodes in the neighbor nodes whose distance between itself and the sink node is smaller than the distance between itself and the sink node. Then the unit distance energy consumption of each node is calculated. After that, the next-hop node is found based on the two-node greedy algorithm. The final optimal path is $A \rightarrow B \rightarrow C \rightarrow D \rightarrow BS$, the total energy consumption is $10e$. If we do not consider the average energy consumption per unit distance, and consider the energy consumption within the two-hop range only, the best path is

A→B→C→I→G→BS, the total energy consumption of $11.5e$. Therefore, the optimal path calculated from the unit energy consumption in the two-hop range approximates the optimal solution approximately.

To simplify the calculation process, a cost factor is defined. Each node chooses the next hop path according to its cost factor:

$$f(i, j, k) = a \frac{e'_{i,j} + e'_{j,k}}{d_i - d_k} + b \frac{\bar{E}_{neighbor}(i)}{E_{current}(j)} + c \frac{e_{non-CH}(j)}{\bar{e}_{non-CH}(i)} \quad (15)$$

where, $f(i, j, k)$ represents the two-hop energy cost estimation factor of node i transmitting data to node k through node j ; $e'_{i,j}$ represents the energy consumed by node i transmitting 100 bit data to node j ; d_i represents the distance from node i to node BS; $\bar{E}_{neighbor}(i)$ represents the average residual energy of the node i 's neighboring cluster head nodes; $E_{current}(j)$ represents the residual energy of the cluster-head j ; $N_{non-CH}(j)$ represents the member nodes number of the cluster-head j ; $\bar{N}_{non-CH}(i)$ represents the member nodes' average number of cluster-heads that are the neighbors of the cluster-head i , and a, b, c satisfies $a+b+c=1$.

3.3 Protocol procedure

The UCEEC algorithm could be divided into three stages: cluster division and clustering, multi-hop path chosen between cluster-heads, data transmission and route update. Each node in the network transmits a Hello message with its maximum transmit power P_{max} . The message includes the value of P_{max} , node ID, and location. Each node establishes a neighbor list according to the Hello message that is received by itself. The node information list is shown in Table 1.

Table 1 Node information list

Field name	Meaning
ID _{neig}	Neighbor node ID
e_{loss}	Path loss in communication
pos	Node location coordinates (centered on the sink node)
QID	Cluster area code (initial 0)
e	Residual energy

After the clustering was completed, each cluster head broadcasts its message CLU_MSG (ID, e , ec) and collected other messages of cluster head, then calculated the distance between the neighbor cluster head nodes, and modified the routing table information.

The main steps are as follows:

- 1) All nodes send 100 bits of data to the sink node with the same power.
- 2) The sink node calculates the cluster information and numbers each cluster area. The node with the highest energy in each cluster area is selected as the cluster head node and the environmental factor of each node is calculated.
- 3) The sink node broadcasts the environmental factor, the cluster ID and the cluster head information of each node.
- 4) Each node obtains a list of neighbor nodes based on the hello message, calculates a list of the forward neighbors.
- 5) Each node chooses the next hop node according to the cost factor.

The innovation of the algorithm is mainly reflected in the following aspects.

- 1) The dense occlusion of crops in wheat field results in the non-uniform channel condition and the non-uniform distribution of environmental parameters. A non-uniform clustering method based on the channel correlation degree between nodes is proposed in this paper, which improves the channel uniformity and data

correlation degree among nodes in the cluster, and can effectively reduce the energy consumption of transmission within the cluster.

- 2) In order to avoid the energy consumption funnel effect of long-distance transmission of remote cluster heads, a multi-hop method between cluster heads based on path energy cost estimation is proposed. The energy consumption of 2-hop path is estimated by the effective displacement in BS node direction, which solves the problem that it is difficult for nodes to obtain the global path loss information. The relay path optimization between cluster heads realizes the energy balance between cluster heads based on local information.

4 Simulation results and analysis

4.1 Simulation data and network setting

In order to verify the performance of this method, protocols such as MREBR (Maximum Residual Energy Based Routing), MEC and SOR were selected to compare with the UCEEC protocol. According to the simulation settings of these comparison protocols^[3,4,10-13,25], the network setting is listed in Table 2.

Table 2 Simulation parameter setting of network

Parameters	Value
Network size	(0,0)-(500, 500) m
Node numbers	1000
The position of sink node	(520 m, 520 m)
Node initial energy	0.5 J
Data frame size	500 bits
Broadcast frame size	100 bits

An experiment of signal propagation was carried in wheat field^[5], the receiving signal is influenced by the distance between transmitter and receiver, antenna height, and the growing crops. Based on these testing data, the simulation data set of this work was established. Considering that the measured results were discrete data. The resolution of the data was 10 m, which cannot well reflect the signal fading between nodes at different distances in farmland. Therefore, this study chose the large-scale model in reference [5] when the antenna height is 100 cm at heading stage to reconstruct the signal fading data, which is near the boundary between large-scale and small-scale. In order to reflect the characteristics of small-scale fading, Perlin noise was added to the reconstruct data. By doing this, the resolution of the data has been raised to 1 m. The RSSI data is shown in Figure 4.

4.2 Parameter selection of clustering region based on dissimilarity

As described in Section 3.1, the number of clusters in the UCEEC protocol was determined by function $\gamma(C)$ and parameter σ . Function $\gamma(C)$ was used to control the extent that the distance between regions was greater than that within regions. That is, the larger the value of $\gamma(C)$ is, the more relaxed the condition of merging regions. The smaller the value of $\gamma(C)$ is, the more strict the condition of merging regions. When the value of $\gamma(C)$ approaches 0, each node can be considered as an independent region. σ is used to prevent the occurrence of long and narrow clustering regions. The larger the σ is, the better the shape of cluster area is and also the more clusters. In order to determine the optimal value of sub regions in the wheat field data set, the number of sub regions with different values of $\gamma(C)$ and σ was compared.

As it is shown in Figure 5, the larger the value of dissimilarity γ between clusters, the fewer the number of clusters. It is proved that the smaller the inter-cluster spacing, the bigger the average intra-cluster similarity. When γ takes more than 75%, the number

of clusters changes slowly, indicating that the clustering program tends to be stable. That is, for the farmland environment, taking the 75% to cluster, the signal attenuation partition caused by farmland environment change tends to be stable. When the value of γ is fixed, as σ increases, the number of clusters increases. Too many or too few clusters are not conducive to data transmission. Therefore, σ takes 0.4 and γ takes 75% in the following simulation.

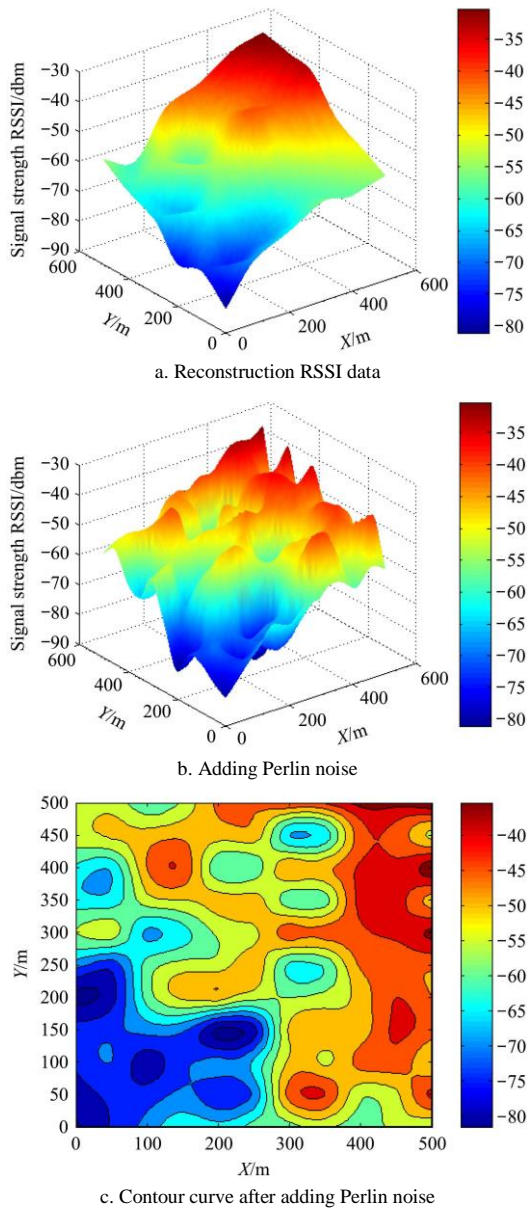


Figure 4 RSSI data in wheat field

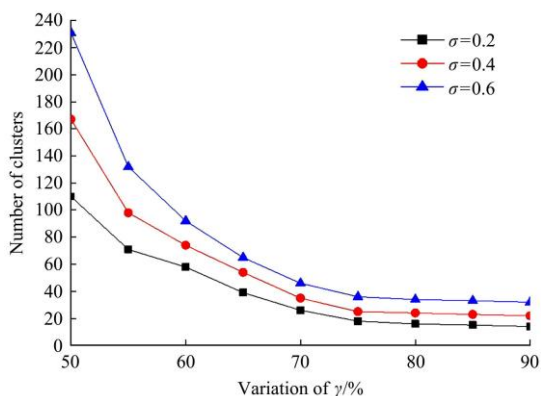


Figure 5 Relation chart of percentages of dissimilarities and the number of clusters

4.3 Simulation results and discussion

During the operation of network, the nodes that consume more energy will die earlier, which will affect the network topology. The more balanced the energy consumption among nodes, the later the death time of the first node appears, and it is beneficial to prolong the lifetime of the network. The slope of the survival node curve also represents the energy consumption efficiency of the network. The larger the slope is, the faster the node dies. It can be considered that the average energy consumption of the network is greater. On the contrary, the smaller slope of the curve, the slower the node death and the smaller the average energy consumption of the network^[22].

As shown in Figure 6, all nodes die at about 320 rounds in MEC, and 340 rounds in SOR. At Round 400, MREBR has about 40% living nodes, while UCEEC has 45%. It can be seen that UCEEC and MREBR have longer network lifecycle than SOR and MEC. The MEC protocol takes the minimum energy path into account without consideration of the node residual energy. Therefore, some nodes consume energy rapidly, and MEC has the shortest network lifetime. Meanwhile, the slope of the UCEEC curve is obviously smaller than the other three protocols. Because the UCEEC algorithm adopts non-uniform clustering based on node similarity, it reduces the average energy consumption within the cluster. It also shows that the UCEEC has a better performance of energy balance, which is partly caused by path selection between cluster-heads.

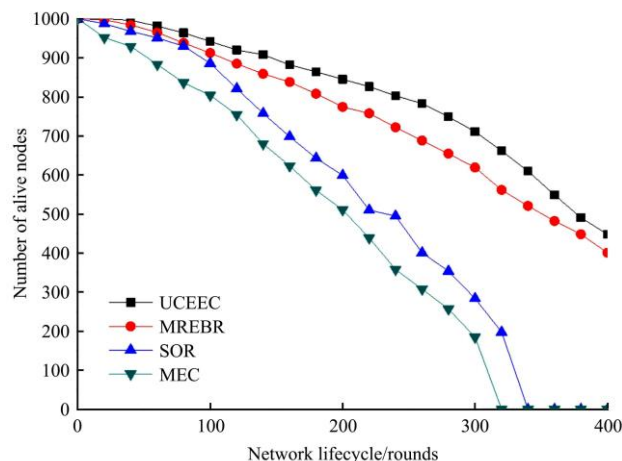


Figure 6 Comparison of different algorithms for network life cycle

As can be seen from Figure 7, when 30% of the nodes die, MREBR takes about 250 rounds, MEC takes about 140 rounds, SOR takes about 160 rounds, UCEEC takes about 300 rounds. The UCEEC protocol have more rounds than the other 3 protocols at the death of first node, 30% nodes, and 70% nodes. It shows that the UCEEC protocol can find out the optimal transmission path that the multi-hop of cluster-heads by considering the impact of environment on signal transmission, reduce the energy loss, and balance the energy efficiency of the whole network, effectively extending the life cycle of the network. The mean value and variance of the residual energy distribution were chosen to illustrate the energy balance performance.

Figure 8 shows the performance comparison of energy balance between the four protocols. In the curve of mean value of the residual energy, a smaller slope indicates a slower energy consumption rate and a longer life time. The slope of UCEEC curve is significantly smaller than that of MREBR, MEC, and SOR, which means that the average energy consumption is lower in the

UCEEC protocol. In the variance curve chart, the MEC curve fluctuates greatly and there are significant peaks around 150 rounds. This indicates that the protocol has uneven distribution of residual energy and is prone to energy voids. When the protocol is terminated, some nodes still have more residual energy. The variance of the UCEEC was very low, indicating that the network node energy could be effectively balanced.

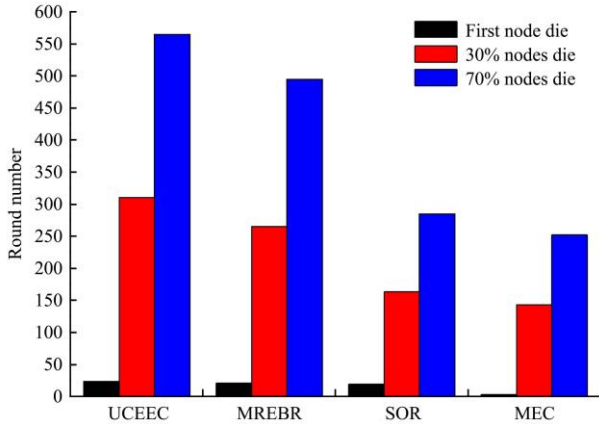
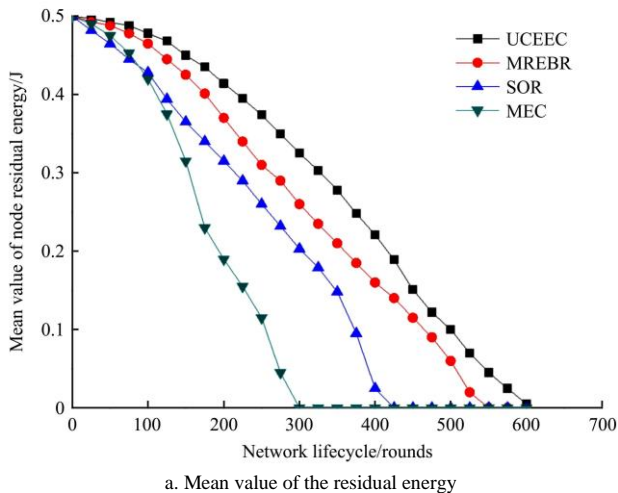
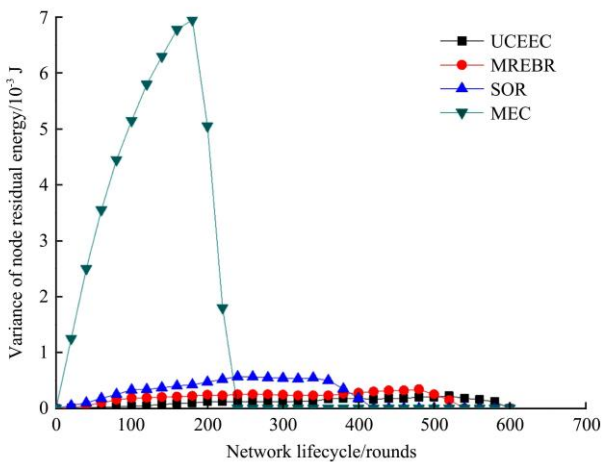


Figure 7 Network rounds of 30% and 70% nodes die



a. Mean value of the residual energy



b. Variance of the residual energy

Figure 8 Comparison of mean and variance of the residual energy

5 Conclusions

Aiming at the requirement of high energy efficiency monitoring in wheat field wireless sensor network, a non-uniform clustering routing protocol based on effective energy consumption is proposed, focusing on the non-uniform multipath fading effect of

crops on wireless signals. This protocol uses the similarity region segmentation between nodes to achieve non-uniform clustering, which improves the similarity of nodes in the cluster in terms of data and channel. A multi-hop path selection method between cluster-heads based on the estimation of two-hop effective energy consumption was designed. The energy consumption cost factor was calculated by the effective energy consumption and the average energy consumption within the cluster to achieve the minimum and balance of the overall energy consumption of the network. The results show that compared with the traditional algorithms MREBR, MEC and SOR, the proposed algorithm improves the energy balance effect between nodes and prolongs the network life cycle. The energy efficient utilization of wireless sensor network data collection in the complex environment of wheat field is realized.

Acknowledgements

The authors acknowledge that this work was supported in part by the National Natural Science Foundation of China (Grant No. 61871041), Beijing Science and Technology Project (Grant No. Z191100004019007), the Youth Found of Beijing Academy of Agriculture and Forestry Sciences (Grant No. QNJ202030), and the Project of Faculty of Agricultural Equipment Department of Jiangsu University (Grant No. 4111680005).

[References]

- [1] Xie C, Zhang D X, Yang L, Cui T, Zhong X J, Li Y H, et al. Remote monitoring system for maize seeding parameters based on android and wireless communication. *Int J Agric & Biol Eng*, 2020; 13(6): 159–165.
- [2] Padilla-Medina J A, Contreras-Medina L M, Gavilan M U, Millan-Almaraz J R, Alvaro J E. Sensors in precision agriculture for the monitoring of plant development and improvement of food production. *Journal of Sensors*, 2019; 2019: 7138720. doi: 10.1155/2019/7138720.
- [3] Fathallah K, Abid M A, Hadjalouane N B. Enhancing energy saving in smart farming through aggregation and partition aware IoT routing protocol. *Sensors*, 2020; 20(10): 2760. doi: 10.3390/s20102760.
- [4] Thakur D, Kumar Y, Kumar A, Singh P K. Applicability of wireless sensor networks in precision agriculture: A review. *Wireless Personal Communications*, 2019; 107(1): 471–512.
- [5] Miao Y S, Wu H R, Li F F, Zhu L. Study of wheat farmland multipath fading channel modeling based on statistical distribution. *Acta Electronica Sinica*, 2016; 44(3): 665–672. (in Chinese)
- [6] Goldsmith A. *Wireless communications*. Oxford City: Cambridge University Press, 2005; pp. 36–46.
- [7] Vougioukas S, Anastassiou H T, Regen C, Zude M. Influence of foliage on radio path losses (PLs) for wireless sensor network (WSN) planning in orchards. *Biosystems Engineering*. 2013; 114(4): 454–465.
- [8] Wang L N, Wang B R. Greenhouse microclimate environment adaptive control based on a wireless sensor network. *Int J Agric & Biol Eng*, 2020; 13(3): 64–69.
- [9] Guo X M, Yang X T, Chen M X, Li M, Wang Y A. A model with leaf area index and apple size parameters for 2.4 GHz radio propagation in apple orchards. *Precision Agriculture*, 2015; 16(2): 180–200.
- [10] Chandrawanshi V S, Tripathi R K, Pachauri R. An intelligent low power consumption routing protocol to extend the lifetime of wireless sensor networks based on fuzzy C-means++ clustering algorithm. *J. Intell. Fuzzy Syst*, 2020; 38(5): 6561–6570.
- [11] Li W, Baoyintu, Jia B, Wang J X, Watanabe T. An energy based dynamic AODV routing protocol in wireless Ad Hoc networks. *CMC-Comput. Mat. Contin*, 2020; 63(1): 353–368.
- [12] Anisi M H, Abdul-Salaam G, Abdullah A H. A survey of wireless sensor network approaches and their energy consumption for monitoring farm fields in precision agriculture. *Precis. Agric.*, 2015; 16(2): 216–238.
- [13] Zhu L, Fan C, Wu H, Wen Z. Coverage optimization strategy for WSN based on energy-aware. *International Journal of Computers Communications & Control*, 2016; 11(6): 877–888.
- [14] Mohamad M M, Kheirabadi M T. Energy efficient opportunistic routing algorithm for underwater sensor network: A review. *2nd ICSITech, IEEE*,

- October 26-27, 2016; pp.41–46.
- [15] Ben Fradj H, Anane R, Bouallegue R. Opportunistic routing protocols in wireless sensor networks. *Wireless Personal Communications*, 2019; 104(3): 921–933.
- [16] Zhang S M, Madadkhani M, Shafieezadeh M A, Mirzaei A. Novel approach to optimize power consumption in orchard WSN: Efficient opportunistic routing. *Wireless Personal Communications*, 2019; 108(3): 1611–1634.
- [17] Alajeely M, Doss R, Ahmad A. Routing protocols in opportunistic networks - A survey. *IETE Tech. Rev.*, 2018; 35(4): 369–387.
- [18] Wu H R, Zhang L H, Miao Y S. The propagation characteristics of radio frequency signals for wireless sensor networks in large-scale farmland. *Wireless Personal Communications*, 2017; 95(4): 1–18.
- [19] Ismail M, Islam M, Ahmad I, Khan F A, Qazi A B, Khan Z H, et al. Reliable path selection and opportunistic routing protocol for underwater wireless sensor networks. *IEEE Access*, 2020; 8: 100346–100364.
- [20] Xuan C Z, Wu P, Zhang L N, Ma Y H, Liu Y Q, Maksim. Compressive sensing in wireless sensor network for poultry acoustic monitoring. *Int J Agric & Biol Eng*, 2017; 10(2): 94–102.
- [21] Chen Y, Shi Y L, Wang Z Y, Huang L. Connectivity of wireless sensor networks for plant growth in greenhouse. *Int J Agric & Biol Eng*, 2016; 9(1): 89–98.
- [22] Pandiyaraju V, Logambigai R, Ganapathy S, Kannan A. An energy efficient routing algorithm for WSNs using intelligent fuzzy rules in precision agriculture. *Wireless Personal Communications*, 2020; 112(1): 243–259.
- [23] Haseeb K, Din I U, Almogren A, Islam N. An energy efficient and secure IoT-based WSN framework: An application to smart agriculture. *Sensors*, 2020; 20(7): 2081. doi: 10.3390/s20072081.
- [24] Khan T H F, Kumar D S. Ambient crop field monitoring for improving context based agricultural by mobile sink in WSN. *Journal of Ambient Intelligence and Humanized Computing*, 2020; 11(4): 1431–1439.
- [25] Liu Y, Tong K F, Wong K K. Reinforcement learning based routing for energy sensitive wireless mesh IoT networks. *Electronic Letters*, 2019; 55(17): 966–968.
- [26] Noh K M, Park J H, Park J S. Data transmission direction based routing algorithm for improving network performance of IoT systems. *Applied Sciences*, 2020; 10(11): 3784. doi: 10.3390/app10113784.
- [27] Varsa G V S, Sridharan D. A balanced energy efficient virtual backbone construction algorithm in wireless sensor networks. *AEU - International Journal of Electronics and Communications*, 2019; 107: 110–124.
- [28] Li D, Li Z. System Analysis and Development Prospect of Unmanned Farming. *Transactions of the CSAM*, 2020; 51(7): 1–12. (in Chinese)
- [29] Chen Y, Shi Y L, Wang Z Y, Huang L. Connectivity of wireless sensor networks for plant growth in greenhouse. *Int J Agric & Biol Eng*, 2016; 9(1): 89–98.
- [30] Liu H, Meng Z J, Wang H, Xu M. Spatio-temporal variation analysis of soil temperature based on wireless sensor network. *Int J Agric & Biol Eng*, 2016; 9(6): 131–138.
- [31] Zhao M, Kumar A, Chong P H J, Lu R. A reliable and energy-efficient opportunistic routing protocol for dense lossy networks. *IEEE Wireless Communication Letters*, 2017; 6(1): 26–29.
- [32] Luo J, Hu J, Wu D, Li R F. Opportunistic routing algorithm for relay node selection in wireless sensor networks. *IEEE Transactions on Industrial Informatics*, 2015; 11(1): 112–121.
- [33] Jawad H M, Jawad A M, Nordin R, Gharghan S K, Abdullah N F, Ismail M, et al. Accurate empirical path-loss model based on particle swarm optimization for wireless sensor networks in smart agriculture. *IEEE Sensors Journal*, 2020; 20(1): 552–561.
- [34] Mansouri M, Ilham O, Snoussi H, Richard C. Adaptive quantized target tracking in wireless sensor networks. *Wireless Networks*, 2011; 17(7): 1625–1639.
- [35] Felzenszwalb P F, Huttenlocher D P. Efficient graph-based image segmentation. *International Journal of Computer Vision*, 2004; 59(2): 167–181.
- [36] Barkunan S R, Bhanumathi V. An efficient deployment of sensor nodes in wireless sensor networks for agricultural field. *Journal of Information Science and Engineering*, 2018; 34(4): 903–918.