Research Article

Chunling Tang*

A clustering algorithm based on nonuniform partition for WSNs

https://doi.org/10.1515/phys-2020-0192
received January 13, 2020; accepted October 01, 2020

Abstract: Wireless sensor networks (WSNs) have great application potential in partition parameter observation, such as forest fire detection. Due to the limited battery capacity of sensor nodes, how to reduce energy consumption is an important technical challenge. In this paper, we propose an energy efficient routing algorithm of adaptive double cluster head (CH) based on nonuniform partition for WSN. Firstly, according to the distance information from the base station (BS) to every sensor node, the network is divided into several uneven partitions. Secondly, CH is selected for each partition as the primary cluster head (PCH). Because of the cluster-level routing, the CHs close to the BS need to forward more data than the CHs in other areas, which consumes more energy. Therefore, an adaptive double CH method can be used to generate a secondary cluster head (SCH) in the cluster near the BS according to the parameters. Finally, the PCH is responsible for data collection, data integration, and data transmission. while the SCH is in charge of data routing. Simulation results show that the proposed algorithm can reduce the energy consumption and extend the life of the WSNs, compared with LEACH protocol and the HEED protocol.

Keywords: WSN, double cluster head, LEACH, nonuniform grid

1 Introduction

Wireless sensor networks (WSNs) are very hot research directions in recent years [1]. This type of network consists of small, inexpensive, and low energy sensor nodes, distributed in specific field to measure and monitor conditions such as forest fire detection. WSNs are powered by batteries, and their energy, communication distance, computing, and storage capabilities are limited. Therefore, how to design efficient routing protocols to prolong the lifetime of the whole network and improve the energy consumption while ensuring the quality of data transmission has become a research hotspot [2]. Clustering routing protocols are the hottest research in routing protocols. Among the clustering routing protocols, the most classical one is the LEACH (low energy adaptive clustering hierarchy) [3–5]. LEACH algorithm is considered more efficient for data transmission than static clustering and direct transmission. However, the algorithm has many drawbacks; the main drawbacks are that the number of clusters is random and there is random election for cluster heads (CHs) without consideration for each node residual energy [6,7]. Random CH selection may cause uneven energy consumption among nodes, thus the network lifetime will be reduced [8,9]. To deal with the problem of random clusters number, the network is divided into several nonuniform partitions; the closer the grid is to the base station (BS), the smaller the grid area is. As for the CHs' energy problem, the adaptive double cluster head method (ADCHM) is used to generate double CHs in the cluster near the BS. The primary cluster head (PCH) is used to collect inter-cluster data. The secondary cluster head (SCH) is used to forward the inter-cluster information and share the communication burden of the PCH. There are four steps in this algorithm. The first step is to divide the observed area into some nonuniform partitions and partition optimization. The second step is to elect the PCH and SCH every partition. The PCH selection is based on the ratio of the residual energy of the node to the average residual energy of the adjacent nodes and the average distance from other nodes in the cluster as the important parameters. The SCH is generated in the partition close to the BS according to some parameters. The third step is inter-cluster routing establishment stage. In this step, the PCH is responsible for collecting data and fusing data in the

* **Corresponding author: Chunling Tang,** College of Electronical Information Engineering, Chongqing Technology and Business Institute, Chongqing 400052, China, e-mail: 2331214397@qq.com

partition, and then transmitting the fusion data to the nearest SCH to BS. The fourth step is the data transmission phase. The minimum spanning tree method is used for data transmission between clusters. The SCH is responsible for forwarding data from the neighbor cluster to the next secondary CHs closer to BS. The SCH is in charge of routing, which forwards the collected data to reduce PCH consumption.

2 The proposed method

2.1 WSN model

WSNs represent a promising solution and have thus gained attention for forest fire detection. Usually, the detection center is located next to the rural road and cannot be moved. Sensor nodes are randomly distributed throughout the forest to detect each area of the forest. Sensor nodes are randomly deployed in a two-dimensional square field; the network model is shown in Figure 1. The features of this network model are as follows:

1. Sensor nodes are distributed in a square area $M \times M$. The network is a stable network with high density and nonuniform distribution. The relationship between grid widths is $w_2 = 1.2 \times w_1$. The relationship between grid heights is $h_2 = 1.2 \times h_1$, $h_3 = 1.2 \times h_2$.

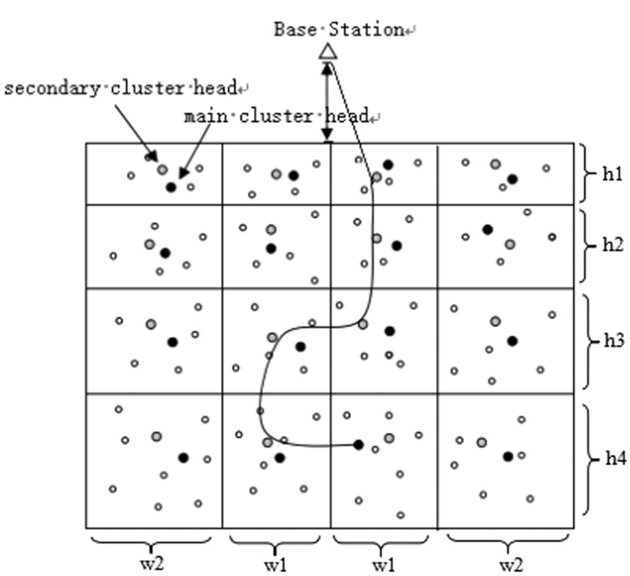

Figure 1: Network model.

2. The sensor nodes have the same initial power and the nodes are non-rechargeable. Sensor nodes are static, each node has a unique identity number, can know its own location, and all nodes can synchronize in time.
3. All nodes can detect their own residual energy and have the function of data collection. Each node works independently and is not affected by other nodes.
4. Each node can computer the distance to another node based on the signal strength of received message and adjust the magnitude of the transmitted power according to the distance.
5. The BS is located at H_m just above the square WSN, and the energy of the BS is infinite.

2.2 Energy consumption model

The simplified energy model of heinzelman [10] is used, which is widely used in various sensor networks. In this model, the free space model and the multipath attenuation model are divided by the communication distance between the source and the destination. In practice, the transmission energy is based on the threshold distance d_t. When the communication distance is less than d_t, the free space model is considered and the parameter ε_{fs} (pJ/(bit m^{-2})) is used. Otherwise, the multipath falloff model and parameter ε_{mp} (pJ/(bit m^{-2})) are used. Sensor nodes are able to process, transmit, and receive data; the node transmits a message with the data communication distance d of k bit; the formulas of used energy of radio model are shown below.

The energy consumption in transmitting data is as:

$$E_T(k, d) = \begin{cases} kE_{elec} + k\varepsilon_{fs} d, & d < d_t \\ kE_{elec} + k\varepsilon_{mp} d^2, & d \geq d_t \end{cases} \tag{1}$$

$$d_t = \varepsilon_{fs} / \sqrt{\varepsilon_{mp}} \tag{2}$$

The energy consumption in receiving data is as:

$$E_R(k, d) = \begin{cases} kE_{elec} + k\zeta_{fs} d, & d < d_t \\ kE_{elec} + k\zeta_{mp} d^2, & d \geq d_t \end{cases} \tag{3}$$

$$d_t = \zeta_{fs} / \sqrt{\zeta_{mp}} \tag{4}$$

In this formulation, k is the number of bytes of data packet, d is distance of transferring [11–13]. When transferring distance is less than threshold value d_t, energy amplification uses free space propagation approach; otherwise, it uses multiroad fading approach. E_{elec} is energy consumed in running electronic circuit in sending and receiving, ε_{fs}, ε_{mp}, ζ_{fs}, ζ_{mp} are, respectively, used energy factors of amplifier circuit in two approaches. The default value is 0.

2.3 LEACH

The steps of LEACH algorithm are as follows:

(1) Each node can be the CH, and each node produces a random numbers rand( ). Compare rand( ) with threshold $T(n)$. If rand( ) ≥ $T(n)$, this node is the CH. The calculation formula of $T(n)$ is as:

$$T(n) = \begin{cases} \dfrac{P}{1 - p(r \mod(1/p))}, & n \in G, \\ 0, & n \notin G \end{cases} \tag{5}$$

When P is the percentage of CH, R is the number of rounds, G is the node set that is not CH.

Once a node becomes a CH, it will broadcast messages to the whole network. Nodes that are not CHs will receive messages broadcast by some CHs, join the cluster where the CH with the largest signal strength is received, and send the join information to the CH. Because the greater the signal strength is, the closer the distance is, and the less energy consumption of receiving and sending messages.

(2) After clustering, the CH generates the time slot table according to the number of nodes in the cluster. Each node can only send data to the CH in its own time slot, otherwise it is in sleep state to save energy.

(3) Data transmission. The CH will turn on the receiver and wait for each cluster member node to send data. The CH will compress and fuse the data and send the data to the BS.

LEACH algorithm uses the random way to determine the CH; the number of members in the cluster is quite different and the load of the CH is unbalanced, which leads to the premature death of some CHs. The node in the cluster uses the single hop way to communicate with the CH, which has a large energy consumption and has a negative impact on the network lifetime.

2.4 Description of the algorithm process

This algorithm uses hierarchical protocol. The network model and energy model of various routing algorithms in WSN are basically designed according to the classical model structure. The concept of “round” is proposed. Each round competes for CH according to the distance between nodes in the cluster and the current energy of nodes [14]. Each round of transmission is divided into four stages: grid initialization stage, election of PCH and SCH stage, inter-cluster routing establishment stage, and data transmission stage. In the cluster, TDMA is used to transmit information, and the routing between clusters is transmitted to the BS node according to the constructed routing tree.

2.4.1 Grid initialization and grid optimization phase

Initially, all nodes in the network are randomly deployed in a square area on the two-dimensional plane, and the BS is set as one side of the square area. When the network deployment is completed, the BS will broadcast a network initialization message. According to the beacon nodes in the network, all nodes in the network calculate their positions and distance from the BS according to the received broadcast signal strength. There are many techniques used to conserve WSN energy, in order to prevent its premature exhaustion [15]. CH near the BS frequently forwards the information sent by other CHs to the BS. To avoid premature consumption of CH near the BS, the area is divided into several grids with unequal width and height. Height increases from near to far equivalence ratio according to distance from the BS, $h_2 = 1.2 \times h_1$, $h_3 = 1.2 \times h_2$, $h_4 = 1.2 \times h_3$... The width increases equally and symmetrically according to the center of the subregion centered on the BS abscissa, $w_2 = 1.2 \times w_1$, $w_3 = 1.2 \times w_2$. CHs near the BS often take on the task of forwarding the data from other CH nodes and tend to die prematurely. Therefore, the grid far from the BS is larger to alleviate the data forwarding task of the nodes near the BS.

When the number of surviving nodes in a grid is less than the set value N_{less}, the grid will be recombined with the adjacent grid to form a new grid. The surviving nodes compete for CHs again in the new grid.

2.4.2 Election of PCH and SCH phase

In WSN clustering routing protocol, the energy consumption of CHs affects the performance of the whole network. When multi-hop routing between clusters is adopted, the CH receives and merges the data. Based on the nonuniform partition adaptive double CH algorithm, we need to determine whether there are double CHs in this partition. The sensor network is divided into areas as shown in Figure 2.

N is the region number, which represents the distance between the node and the BS. The larger the n is, the farther it is from the BS. Multi-hop communication leads to more forwarding tasks and more energy consumption for CHs in the area closer to the BS. If closer to the BS, the more CHs can undertake forwarding tasks, that is, the greater the distribution density of CHs, the less energy each CH consumes in forwarding data. In order to increase the distribution density of CHs in the area closer to the BS, the probability of SCH generation in each region is set to $P = 1/n^3$. Considering the residual energy of nodes, the competitive parameters of SCH generation are obtained as follows:

$$T_{\mathrm{SCH}} = \frac{E_{\mathrm{residual}}}{E_{\max}} + \frac{1}{n^3} \tag{6}$$

According to the observation area, the competition parameter P_{SCH} is set to compare with T_{SCH}. If T_{SCH} is larger than P_{SCH}, the cluster has two CHs, otherwise the cluster has one CH.

At the beginning of each round, all nodes in WSN participate in the competition of CH election, and then the final CH is determined by time sequence. According to the principle of selecting the PCH, SCH, and the data collection, it is necessary to ensure that any two nodes in the two adjacent subregions in the longitudinal direction can reach one hop. Therefore, the communication radius R of nodes should be larger than the farthest distance that two nodes in adjacent layers. The Communication distance R is as follows:

$$R_{\mathrm{t}} \geq \sqrt{(d_{\mathrm{H1}} + d_{\mathrm{H2}})^2 + (d_{\mathrm{Z1}} + d_{\mathrm{Z2}})^2} \tag{7}$$

In this formulation, d_{H1}, d_{H2} is the width of the H1-th, H2-th subregion, d_{Z1}, d_{Z2} is the height of the Z1-th, Z2-th subregion. Therefore, if the initial communication radius of the primary and secondary CHs is less than R_{t}, the transmission power needs to be increased. Considering the shortcomings of CH competition in LEACH, the ratio of residual energy of nodes in the same grid to average residual energy of neighboring nodes and the average distance between nodes and other nodes in the cluster is taken as an important parameter to participate in the competition of CHs.

In the first round, the nodes nearest to the grid center are the PCHs and the SCHs. After the first round of data collection, the residual energy of each node will be different. In this case, the CH competition parameter $s(n)$ is as:

$$s(n) = \begin{cases} \alpha\dfrac{E_{\mathrm{current}}}{E_{\mathrm{ave}}} + \dfrac{\beta D_{\mathrm{ave}}}{M}. & n \in G \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

When the residual energy of node n is larger than the average energy of its neighbors, the CH competition parameters are calculated according to its residual energy and the number of rounds it has been elected; where α, β is a constant coefficient, E_{current} is the current energy of node n, E_{ave} is the average energy of neighbor nodes, M is the number of rounds, and D_{ave} is the average distance between nodes and other nodes in the cluster. When the residual energy of node n is less than the average energy of its neighbors, the node does not participate in CH competition, and its competition parameter is 0. The nodes in each grid are sorted from large to small according to their competition parameters. If a region has only one CH, the node with the largest energy is the PCH. If a region has two CHs, then they are the PCH and the SCH.

Figure 2: Network area division diagram.

2.4.3 Inter-cluster routing establishment

Since this round of grid has been determined, the broadcast radius of each grid is set to d. In order to enable the adjacent SCHs to communicate with each other, all SCHs broadcast the SCH routing message with the broadcast radius of $2.5 \times d$, which includes the ID of the SCH, the number of hops from the BS, and the remaining energy of the SCH. After receiving the broadcast message, the SCHs record the relevant information of the communication CH to their own communication routing table. When all SCHs broadcast the routing message, all SCHs will get the information about their neighbor communication CHs. In order to improve the energy efficiency and balance the network load, the SCH will construct multi-hop routing according to the information in its communication routing table. Among all the neighbor SCHs, it is necessary to select an optimal communication SCH as the next hop routing node. To select the optimal SCH, we need to consider the factors such as the smaller distance from the BS and the more residual energy. If the BS is within the communication radius of the SCH, the BS will be the next hop node directly. According to the residual energy and signal strength of each neighbor SCH in the communication CH routing table, the SCH will select the optimal communication CH as the next hop routing node.

2.4.4 Data transmission phase

In this phase, the minimum spanning tree algorithm is used for inter-cluster routing. PCH is responsible for sending the data collected in its grid to the next hop SCH. SCH is responsible for forwarding the data between grids.

The data transmission is divided into several time slots, each of which includes two parts: the data set in the cluster and the data transmission of the communication CH. In this way, it can reduce the channel conflict and the energy consumption of nodes, so as to extend the network life cycle.

In the data collection phase, each cluster member node can only collect data in its own time slot and send it to the PCH of its own cluster in a single hop way. In the time slot that does not belong to itself, the member nodes of the cluster will remain dormant to save energy. After PCH collects the data of cluster members, the collected data are fused. In order to reduce the amount of data transferred and the energy consumption of CH node, the CH node sends the fused data to its neighbor SCH. The SCH is responsible for data routing. If at least one of the PCH and SCH needs to be updated, it needs to be broadcasted in the grid. After data fusion, PCH transfers data to BS in multi-hop way. When the distance between the CH and BS is less than the communication radius, the CH directly sends the data to BS without forwarding, otherwise it can only transmit the data to BS in the way of multi-hop. Obviously, the SCH should meet the following relationship:

$$\{\mathrm{CH}_n | d(\mathrm{CH}_n,\mathrm{BS}) < d(\mathrm{CH}_m,\mathrm{BS}) \;\&\; d(\mathrm{CH}_m,\mathrm{BS}) > d(\mathrm{CH}_n, \mathrm{CH}_m)\} \tag{9}$$

Taking two hop routing as an example, suppose that the data is sent from PCH CH_j and forwarded to BS through routing node CH_i, and the total energy consumed $E_{j\text{-BS}}$ is:

$$E_{n\text{-BS}} = E_{n-m\text{-Tx}}(l) + E_{m\text{-Rx}}(l) + E_{m\text{-BS-Tx}}(l) \tag{10}$$

$E_{n-m\text{-Tx}}(l)$ is the energy consumed by CH_n to send l-bit data to CH_m. $E_{m\text{-Rx}}(l)$ is the energy consumed by CH_m to receive l-bit data. $E_{m\text{-BS-Tx}}(l)$ is the energy consumed by CH_m to send l-bit data to BS.

$$\begin{aligned} E_{n\text{-BS}} &= E_{n-m\text{-Tx}}(l) + E_{m\text{-Rx}}(l) + E_{m\text{-BS-Tx}}(l) \\ &= lE_{\mathrm{elec}} + l\varepsilon_{\mathrm{fs}} d_{n-m}^2 \\ &\quad + lE_{\mathrm{elec}} + lE_{\mathrm{elec}} + l\varepsilon_{\mathrm{fs}} d_{m\text{-BS}}^2 \\ &= 3lE_{\mathrm{elec}} + l\varepsilon_{\mathrm{fs}}(d_{n-m}^2 + d_{m\text{-BS}}^2) \end{aligned} \tag{11}$$

It can be seen that the energy consumed by this data relay transmission is mainly determined by $(d_{n-m}^2 + d_{m\text{-BS}}^2)$. Therefore, when selecting a routing node, the SCH with the smallest $(d_{n-m}^2 + d_{m\text{-BS}}^2)$ should be selected as the forwarding node.

3 Simulation and result

In order to verify the performance of the algorithm of ADCHM, a WSN simulation environment is established in MATLAB environment. Hundred sensor nodes are randomly distributed in the 100 m × 100 m area. The BS is located at (50, 150) and the sampling period is 10 s. After all the nodes are laid out, the position will not change. Without considering the external destructive factors, when the energy of the node is 0, the node will fail. The lifetime of the network is the time of the first node die. The proposed method is compared with LEACH and HEED. The network lifetime in the network is used as the research focus for the experimental simulation. The parameters are shown in Table 1.

The network lifecycle simulation is shown in Figure 3. It can be seen from the table that the proposed method is the lowest and the network lifetime is obviously prolonged.

Table 1: Simulation setting

Parameter	Value
E_0	0.5 J
E_{DA}	5 nJ/bit
E_{elec}	25 nJ/bit
ε_{fs}	10 pJ/(bit m^2)
ε_{mp}	0.0013 pJ/(bit m^2)
H_m	50 m
L_m	100 m
n	100

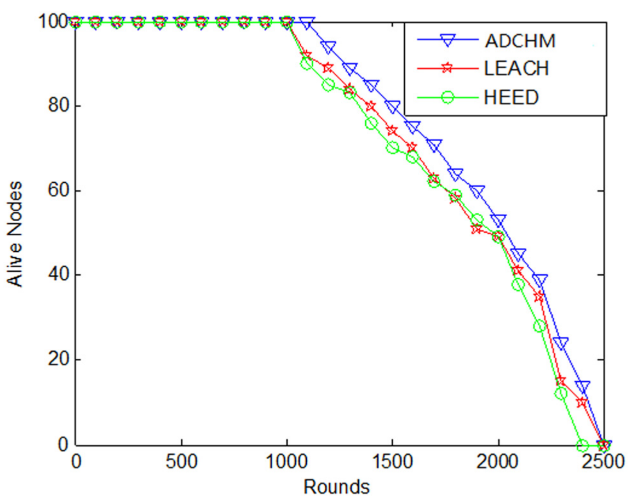

Figure 3: Comparisons of simulation performance between the ADCHM algorithm and LEACH and HEED algorithms.

4 Conclusion

In this paper, based on the analysis of LEACH protocol, an adaptive double CH based on nonuniform grid algorithm is proposed to solve the problems of the random number of clusters and the residual energy of the CHs elected. The network is divided into several uneven grids, which not only avoids the problem of too large or too small clusters, but also avoids the consumption of a CH too fast. Considering the residual energy of nodes and their neighbors when choosing CHs can protect the nodes whose residual energy is lower than the average residual energy of neighbors from becoming hot spots. This method prolongs the lifetime of the whole network. At the same time, in the process of CH election, the way of choosing the one CH or the two CHs enhances the efficiency and robustness of the whole network. Nonuniform partitioning network protects CHs near the BS and avoids them not only aggregate other node data in the region, but also transmit the data to the BS, so that the energy consumption in the network is more balanced. The simulation results show that this method balances the energy consumption of each node in the network and prolongs the lifetime of the network. In the future work, this algorithm will be further studied in the irregular WSN.

Acknowledgments: This research was supported by Scientific and Technological Research Program of Chongqing Municipal Education Commission (No. KJZD-K201804001), and Research Foundation of Chongqing Radio & TV University (Project No. ZZ2020-06), and partly funded by Natural Science Foundation of Chongqing (Grant No. cstc2020jcyj-msxmX0196) and Scientific & Technological Research Program of Chongqing Municipal Education Commission (Grant No. KJQN202004001).

Reference

[1] Morsy NA, Abdelhay EH, Kishk SS. Proposed energy efficient algorithm for clustering and routing in WSN. Wirel Personal Commun. 2018;103(1):2575–98.

[2] Zheng ZY, Guo F, Wang ZF, Zhang XJ, Wang F. Research on routing protocol based on gradient and energy awareness in wireless sensor networks. Computer Sci. 2017;44(2):115–9.

[3] Heinzelman WB, Chandrakasan AP, Balakrishnan H. An application-specific protocol architecture for wireless microsensor networks. IEEE Trans Wirel Commun. 2002;1(4):660–70.

[4] Peng L, Lu JX, Liu QP. Research on hybrid LRACH protocol for large scale wireless sensor networks. Chin J Sens Actuators. 2016;29(11):1737–41.

[5] Younis O, Fahmy S. HEED: A hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. IEEE Trans Mob Comput. 2004;3(4):366–79.

[6] Dutta R, Gupta SK, Das M. Improvement on LEACH protocol in wireless sensor networks. Int J Comput Appl. 2014;97(21):36–40.

[7] Saini P, Sharma AK. Energy efficient scheme for clustering protocol prolonging the lifetime of heterogeneous wireless sensor networks. Int J Comput Appl. 2010;6(2):30–6.

[8] Awwad SAB, Ng CK, Noordin NK, Rasid MFA. Cluster Based Routing Protocol for Mobile Nodes in Wireless Sensor Network. Springer US; 2011. p. 251–81.

[9] Kuila P, Jana PK. A novel differential evolution based clustering algorithm for wireless sensor networks. Appl Soft Comput. 2014;25:414–25.

[10] Donoho D. Compressed sensing. IEEE Trans Inf Theory. 2006;52(4):1289–306.

[11] Butanda JA, Málaga C, Plaza RG. On the stabilizing effect of chemotaxis on bacterial aggregation patterns. Appl Math Nonlin Sci. 2017;2(1):157–72.

[12] Esteban M, Enrique PN, Torres F. Bifurcation analysis of hysteretic systems with saddle dynamics. Appl Math Nonlin Sci. 2017;2(2):449–64.
[13] Gao W, Wang WF. The fifth geometric-arithmetic index of bridge graph and carbon nanocones. J Differ Equ Appl. 2017;23(1):100–9.
[14] Mi C, Shen Y, Mi WJ, Huang YF. Ship identification algorithm based on 3D point cloud for automated ship loaders. J Coast Res. 2015;73(1):28–34.
[15] Sudhakar S, Francis S, Balaji V. Odd mean labeling for two star graph. Appl Math Nonlin Sci. 2017;2(1):195–200.