

A New WSN Clustering Routing Algorithm with Energy Consumption Balance

Zhongxiang XIAO & Xiaoli NIE

School of Electronic Engineering, Xi'an Shiyou University, Xi'an, China

Xiaoying YAN, Guojian CHENG & Xinjian QIANG

School of Computer Science, Xi'an Shiyou University, Xi'an, China

ABSTRACT: Based on Leach-C, this paper proposes a novel routing algorithm for energy consumption balance, named EBLEach-C. Using Self-organization Mapping and Kmeans clustering algorithm, the nodes that are adjacent with equal energy are self-organized into a cluster. A new cost function is introduced to select an optimum relay node between cluster head and base station. Simulation results demonstrate that the proposed algorithm can effectively avoid the early dying of cluster which is far away from base station, and it can also balance the energy consumption, and improve the network coverage range.

KEYWORD: Wireless Sensor Network; Self-organization Mapping; Clustering routing algorithm; Network life cycle; Particle Swarm Optimization

1 INTRODUCTION

Hierarchical routing based on the clustering is the most efficient routing algorithm in wireless sensor networks (WSN), because it is based on the thought of "divide and conquer". After clustering, cluster head (CH) collects and integrates of data from the normal nodes, and send the data to the base station (BS). Hierarchical communication reduces the conflict among channels, and saves energy consumption and prolongs the network life cycle. Leach and Leach-C are first used for WSN hierarchical routing [1], most hierarchical routing algorithms are developed on this basis. Currently the improvements for WSN clustering routing algorithm are generally about the formation of the cluster, CH selection and communication mode. Because of the defect of Leach algorithm, EBLP agreement is proposed [2], by controlling the number of members in the cluster and forming a tree of routing chain among the cluster nodes and CH nodes to save energy consumption of the cluster head. The WSN asymmetric centralized energy efficient routing algorithm [3] based on Leach-C is proposed to form a shortest multi-hop data transmission route, by reverse seeking shortest path to base station. However, the algorithm above only based on the coordinates of nodes to select CH node, so energy consuming and transit delaying are large during the multi-hop routing.

WSN and artificial neural network structure are quite similar. The nodes with the ability of

perceiving and processing in WSN can be seen as neurons, connection between nodes corresponds to the connection of synapses for signal transmission. Self-organizing mapping (SOM), neural network, Particle Swarm Optimization (PSO) and other intelligent algorithms have been used in many references for routing discovery, node clustering, and data fusion in wireless sensor network [4]. References [5, 6] used PSO algorithm, on the basis of Leach, to optimize the cluster head selection with the quantity of node energy and location information, but the arbitrary nature of the Leach algorithm for cluster heads selection results in uneven cluster size and location, thus affecting the overall performance of the algorithm. Lea2c algorithm [7] used SOM for node clustering. Based on Leach-C algorithm framework, this paper proposed SOM+Kmeans clustering method at base station for self-organizing clusters, according to the energy consumption and the network topology, which is referred to as energy balanced consumption routing algorithm, EBLEach-C.

2 SELF-ORGANIZING FEATURE MAPPING ALGORITHM

SOM is an unsupervised neural network structure, the main objective is to translate input signal pattern of arbitrary dimension into a one-dimensional or two-dimensional discrete mapping, and adapt this transformation in topology and orderly manner [8],

SOM algorithm can be described as follows:

Step1: Initialization: select a random value as the initial weight vector $\omega_i(0), i = 1, 2, \dots, k$ (k is the number of neurons), generally differ from each other, which can be randomly selected from the input vector weights, with ω_i the same dimension as input vector.

Step2: Load sample: at time step t , the input pattern vector x is loaded into the network. x can be selected at random, and can also be sequentially input from the training data set.

Step3: Find the winning element: using the minimum Euclidean distance criterion to find the neurons with smallest distance:

$$\|x(t) - \omega(t)\| = \min_i \{\|x(t) - \omega_i(t)\|\}, i = 1, 2, \dots, k \quad (1)$$

where, $c = \arg \min_i \{\|x(t) - \omega_i(t)\|\}, i = 1, 2, \dots, k$, k is the winning neuron index.

Step4: Update the weights: use the updating formula to adjust the right value of the winning neurons c and its neighborhood $N_c(t)$ in other neurons.

$$\omega_i(t+1) = \begin{cases} \omega_i(t) + \eta(t)h_{ci}(x(t) - \omega_i(t)) & i \in N_c(t) \\ \omega_i(t) & i \notin N_c(t) \end{cases} \quad (2)$$

where, $\eta(t)$ is a learning rate parameter; h_{ci} is a topological neighborhood function around winning neurons, usually $\eta(t)$ and h_{ci} diminishing in the process of learning. According to experience, the learning rate is set by: $\eta(t) = \eta_0 \exp(-t/\tau_1)$, $\eta_0 = 0.1$, $\tau_1 = 1000$, $0.01 < \eta(t) < 0.1$, h_{ci} is the Gauss neighborhood function, $h_{ci}(t) = \exp(-d_{c,i}^2 / 2\sigma^2(t))$, h is shrink over time, this requires the width σ decreased with time to meet, i.e. $\sigma(t) = \sigma_0 \exp(-t/\tau_2)$ $t = 1, 2, \dots$, the σ_0 equal to the grid radius, τ_2 is another time constant values: $\tau_2 = 1000 / 1b\sigma_0$.

Step5: Repeat Step2 to select a new input vector, until reaching to the maximum number of iterations t_{max} .

3 WSN CLUSTERING ROUTING ALGORITHM

Compared with the Leach-C algorithms, the SOM algorithm organizes clusters itself, based on the coordinates of the nodes and the residual energy. The proposed algorithm appropriately extends the time for each round of the communication phase, specifically, after the node CH sending packets within a cluster, the cluster head poll to select the new cluster head node with maximum energy, until

the energy of one CH reduce to 3/5 of the energy, it requests base station to regenerate the clusters.

3.1 Cluster building phase

Self-organizing clustering algorithm uses SOM+Kmeans clustering method [9]. Firstly, SOM algorithm is used to preprocess the raw data to obtain the number of clusters and the initial cluster centers and other information, and then based on the results after the first stage. Kmeans algorithm is used to create the final node cluster. The first stage cluster is based on the location and energy of the nodes. The input data set of SOM clustering structure is $[X, Y, E]$, where, X, Y is nodes coordinates; E is the node energy available. In the second stage, cluster number K and cluster centers $C = (c_1, c_2, \dots, c_k)$ are the input parameters in Kmeans algorithm. After iterating many times, it is convergence and output clustering information, completing node clustering of wireless transmission sensor network.

3.2 Cluster head selection phase

The self-organizing clustering result of coordinates of nodes and energy is that the similar characteristics nodes are considered as the same type. Therefore, the nodes in the cluster are adjacent, and the energy distribution is uniform. By references [7], the nodes of maximum residual energy are selected as cluster head.

3.3 Communication phase

After the establishment phase, it is the communication phase of long duration in which the node energy consumption is proportional to transmission distance [1]. Clustering scheme above produces adjacent nodes and evenly distributed energy clusters. Single-hop path selection model will result in premature death of the far from the base station cluster, bringing bad influence in network coverage of the WSN. In the multi-hop communication system, the relay node for receiving and forwarding data are energy consumption, which needs to maintain a routing table of each node.

To overcome the above-mentioned problems, the cost function is proposed f_{cost} , and the largest $f_{cost}(i)$ cluster head is as the best relay node between the current CH and BS, that is, cluster head far from the base station can be two-hop to make the fusion data transmission to the base station through other cluster heads, to avoid the phenomenon of clusters away from the BS premature death.

The relay node is selected by using cost function Equation 3, the distance and residual energy between candidate cluster head and base station are given:

$$f_{cost}(i) = \varepsilon_1 \frac{1}{D_{total}(i)} + \varepsilon_2 \frac{E_i}{E_{avg}} \quad (3)$$

where $D_{total}(i)$ is the sum of the distance from current CH to BS through CH_i , that is $dist(CH2CH_i) + dist(CH_iBS)$. In this paper, the network area and base station close to its cluster head, using single hop transmission mode. $E(i)$ is currently available energy of candidate cluster head CH_i , E_{avg} is average energy of cluster head, $\frac{1}{D_{total}(i)}$ is reciprocal for distance, its maximum value ensure that the selected relay node is the closer solution to base station. $\frac{E(i)}{E_{avg}}$ ensures that the selected relay node energy is bigger, ε_1 and ε_2 is constant, and $\varepsilon_1 + \varepsilon_2 = 1$, setting ε_1 and ε_2 in different values can adjust the proportion of distance and energy factors in selecting relay node, select the CH_i which makes $f_{cost}(i)$ is biggest as a relay node point, determine the ε_1 and ε_2 is 0.3 and 0.7 respectively after many simulation experiments.

3.4 EBLEach-C algorithm

The EBLEach-C algorithm is described as follows:

Step1: Network deployment: given a topological space randomly generate N isomorphic sensor nodes, with the same energy.

Step2: Establishment stage:

Step2.1 Send its node position and energy information to the base station.

Step2.2 Base station generates a adjacent position by using SOM+Kmeans algorithm, energy is clustered uniformly, select the maximum energy of nodes in cluster CH (If the energy is same choose one randomly).

Step2.3 Base station node types and relations (CH or common node belong to which cluster): Code Division Multiple Access (CDMA) expansion frequency code information, Time Division Multiple Access (TDMA) time slot information, nodes receiving and save (for CH nodes, broadcast its coordinates and energy information, record all CH information at the same time).

Step2.4 CH calculation and maintenance of two vectors, one is the distance and energy to other CH , another is distance from all CH to BS .

Step3: Communication:

Step3.1 CH chooses maximum nodes value of, and transmits it to the base station.

Step3.2 If CH energy reduced to 3/5 times of the clustering energy, go to Step2, else Step3.3.

Step3.3 CH poll, choose the node whose residual energy is maximum in cluster as the new CH , go to Step3.1.

4 ALGORITHM SIMULATION AND ANALYSIS

To verify the performance of the algorithm, we use simulation tools Matlab7.0 for experiment testing, the main parameters: network model area is $200m \times 200m$, 200 nodes are randomly distributed in this regions, a base station is located in coordinates (250, 275) place. Before each round of the cluster head, nodes in the cluster is sent to 4000 bit data packets to CH , the data packet size remains 4000 bit after fusion in CH , the data reported to the base package for 100 bit.

4.1 Network model

This algorithm assumes that BS has sufficient energy supply to finish the operations with high energy consumption, such as clustering, cluster head selection and route calculation. The characteristics of network model are as follows: (1) N sensor nodes are randomly distributed in a uniform area A (choose rectangular region), the base station is located outside of the more distant from the region, sensor nodes and base station are stationary. (2) The function is equivalent for each nodes, both have the ability of data sensing and fusion, but node energy is limited. (3) Each node can perceive their location and inform the coordinate information to base station. (4) Nodes perceive the environment with fixed frequency, and the data is always sending. (5) Nodes can perceive their own residual energy and change the transmitting power.

4.2 Energy model

Using the energy consumption model proposed in references [1-6], that is the initial energy of each node is 0.5J, sends k bit data to the node with distance d , needs to consume energy:

$$E_{TX}(k, d) = \begin{cases} k \cdot E_{elec} + k \cdot \varepsilon_{fs} \cdot d^2 & \text{if } d < d_0 \\ k \cdot E_{elec} + k \cdot \varepsilon_{mp} \cdot d^4 & \text{if } d > d_0 \end{cases} \quad (4)$$

The energy consumption of the circuit to receive k bit data from other nodes is as shown below:

$$E_{TX}(k) = k \cdot E_{elec} \quad (5)$$

where E_{elec} is circuit energy consumption for sending and receiving 1 bit data, ε_{fs} and ε_{mp} are constants related to transmission channel model. E_{elec} is specific set to $50 nJ/bit$, and ε_{fs} to

10 pJ/b/m², ϵ_{mp} for 0.0013 pJ/b/m⁴. During the cluster head and the relay node for local processing and data fusion, the energy consumption ϵ_{da} per 1 bit data is $5nJ/b$, with distance constant

$$d_0 = \sqrt{\frac{\epsilon_{fs}}{\epsilon_{mp}}}$$

4.3 Analysis of Results

EBLeach-C and *Leach-C*, *Lea2C* algorithm are compared in simulated results to show the performance of the algorithms. The network life cycle is the evaluation index, death with first nodes, 20% nodes and the last node death rounds to characterize. In the case of *EBLeach-C* algorithm, network energy distributed in the network nodes uniformly, the nodes far away from the base station will always survive and collect information. There are less than 30% nodes alive in *Leach-C* algorithm when the first death of nodes appears in *EBLeach-C* algorithm. Compared with *Lea2C*[7], the death rounds of first node is 901 in *EBLeach-C* algorithm; while 700 rounds in *Lea2C*, delay of nearly 15%, the death time of 20% nodes in *EBLeach-C* algorithm are much later than that in *Lea2c* algorithm. The reason is that in the *Lea2C* algorithm, for a period of time after the operation of the network, the uneven of *CH* energy consumption lead to earlier death of *CH* which is far away from BS. As is shown that *Lea2C* algorithm has a longer network life cycle, since that the energy consumption of partial nodes is less than others' caused by the uneven of energy consumption of this algorithm, the node lifetime will be relatively long, but the network monitoring is meaningless when most of the nodes in the network are death, network life cycle in essence equivalent to the end.

5 CONCLUSION

Limited energy cannot be supplemented twice in

WSN nodes, and efficient network routing algorithm should be designed to prolong the lifetime of network. This paper presents a new WSN clustering routing algorithm based on energy and topology information. The SOM+Kmeans algorithm is applied to clustering node, and the best relay node between cluster head and base station is selected by using a cost function. The effectiveness of the algorithm is demonstrated by the results of simulation. The other version and parameter optimization of SOM algorithm can be further studied.

REFERENCES

- [1] Heinzelman W, Chandrakasan A & Balakrishnan H. 2000. Energy-efficient Communication Routing Protocols for Wireless Microsensor Networks. *Proc. of the 33rd Hawaii International Conference on System Sciences* 8:3005-3014. USA: IEEE Computer Society.
- [2] He Yanjie & Li Layuan. 2009. Energy balanced clustering routing protocol design in WSN. *Journal of transducer technology* 22 (10): 1500-1514.
- [3] Zhu Zijian & Zhao Guangshe. 2009. Asymmetric centralized energy efficient routing algorithm of WSN. *Computer Engineering* 35 (3): 155-157.
- [4] Enami N, Moghadam R A & Dadashtabar K. 2010. Neural Network Based Energy Efficiency in Wireless Sensor Networks: A Survey. *International Journal of Computer Science & Engineering Survey* 1 (1): 39-53.
- [5] Liang Ying & Yu Haibin. 2006. Optimization of routing protocols for WSN based on the clustering applying PSO. *Control and Decision* 21 (4): 453-456.
- [6] Han Dongxue & Zhang Ruihua. 2010. Dual cluster head clustering algorithm of Wireless sensor network based on PSO. *Computer Engineering* 36 (10): 100-102.
- [7] Dehni L, Kief F & Bennani Y. 2005. Power Control and Clustering in Wireless Sensor Networks. *Proceedings of IFIP Med-Hoc-Net 2005. Mediterranean Ad Hoc Networking Conference*: 1-11.
- [8] Cheng Guojian. 2008. The neural network and the growing self-organization network. *Journal of Xi'an Jiaotong University*: 113-135.
- [9] Vesanto J & Alhoniemi E. 2000. Clustering of Self-organizing Map. *IEEE Transactions on Neural Networks* 11 (3): 586-600.