

Distributed Fuzzy Sets-based Clustering Algorithm for Heterogeneous Wireless Sensor Networks

Shao-Juan Xu^{1,a}, Hai-Xia Yu¹, Li-Qiang Diao¹, Zhong-Gao Sun²

¹City Institute, Dalian University of Technology, Dalian, 116600, China

²School of Physics and Electronic Technology, Liaoning Normal University, Dalian, 116029, China

^aemail: bluess18@sina.com

Keywords: Wireless Sensor Networks; Clustering; Energy Heterogeneous; Network Lifetime

Abstract. Clustering provides an effective way for data gathering in wireless sensor networks. In this paper, a new distributed clustering algorithm called DFSC that maximizes lifetime for energy heterogeneous wireless sensor networks is proposed, in which the cluster head selection approach is based on the concept of fuzzy sets. In DFSC, a membership function, a fuzzy enhancement function, and a converting function are interwoven during the cluster head selection phase. As a result, a node with higher residual energy and lower communication cost to base station will have more chance to become a cluster head. So DFSC can better handle the heterogeneous energy capacities. Simulation results under various network scenarios show that this algorithm outperforms some existing clustering methods, in terms of both the network lifetime as well as the network data capacity.

Introduction

Wireless sensor networks (WSNs) consist of numerous low-cost and low-power sensor nodes which have limited processing, storage and radio capabilities [1]. Since the sensor nodes are often powered by batteries which may be difficult or even impossible to replace and recharge, energy efficiency has become the major concern in the design of WSNs.

Clustering is an energy-efficient method used to manage energy consumption for prolonging the network lifetime. In order to prolong the lifetime of a WSN, many clustering algorithms [2-10] have been proposed. Most of the clustering algorithms focus on the homogenous sensor networks. In fact, it is impossible for every node to use its own energy uniformly. Therefore, sensor networks present multilevel energy heterogeneous.

In [3] and [4], the Stable Election Protocol (SEP) algorithm and Distributed Energy-Efficient Clustering algorithm (DEEC) extend the famous clustering algorithm LEACH to multilevel energy heterogeneous networks. Backoff Strategy clustering scheme [5] proposes a novel time-driven cluster head (CH) selection algorithm. Every node generates a waiting time according to an exponential random distribution. A node with higher residual energy is more probable to become a CH.

In this paper, a new distributed clustering scheme, called Distributed Fuzzy Sets-based Clustering (DFSC) algorithm is proposed for multilevel energy heterogeneous WSNs. We first get the membership of the node through fuzzy enhancement algorithm according to its energy and distant to the base station (BS). And CH selection is primarily based on the membership of a node and its neighbor nodes. Thus an energy-rich and lower communication cost node around the weak nodes will produce a shorter waiting time for CH competition. Simulation results show that DFSC can balance the energy load and increase the lifetime of the network significantly.

Detail Scheme of the DFSC Algorithm

In this section, we describe the operational details of the DFSC scheme, whose goal is to increase the lifetime and stability of the network in the presence of heterogeneous nodes. We assume that the initial energy of nodes is uniformly distributed over the close set $[E_{\min}, E_{\max}]$,

where E_{\min} is the lower bound and E_{\max} is the upper bound. And each node is aware of the distribution of initial energy before being deployed. We use similar radio model as proposed in [2]. DFSC works in rounds as LEACH and each round comprises CH selection phase, cluster formation phase and data transmission phase.

A. Cluster Head Selection

The CH selection phase consists of the following steps:

Step1: The first step in this phase is neighborhood discovery. In this step, each node broadcasts a “hello” message using the carrier-sense multiple access with collision avoidance (CSMA/CA) MAC protocol. The node sends this small message containing its ID, distant to the BS and remaining energy to its neighbors within the communication radius r_c . If a node receives a hello message from its n neighbor nodes it will save the information in the neighbor list $nbrlist = \{ID, e_i, dtoBS_i\}, i = 1, \dots, n$.

Step2: After neighborhood discovery completed, node i gets its membership through the membership function

$$\mu_i = G(m_i) = \left[1 + \frac{M_{\max} - m_i}{F_d} \right]^{-F_e} \quad (3)$$

where $M_{\max} = e_{\max} / dtoBS_{\min}$, $m_i = e_i / dtoBS_i$, and the positive constant F_d and F_e are the denominational and exponential fuzzy generators. In general, $F_e = 2$.

The membership function G is proposed by S. K. Pal in [11], which is a popular fuzzy enhancement algorithm used in pattern recognition and image processing.

Using (3), the set $M = \left\{ \frac{e_i}{dtoBS_i}, i = 1, \dots, n \right\}$ is mapped to the fuzzy property plane $U = \{\mu_i, i = 1, \dots, n\}$. The value of μ_i denote the membership degree of m_i towards the ratio of energy and distant to the BS M_{\max} . Obviously the membership function $G(m_i)$ is monotonically increasing and the maximum value of μ_i is 1. It is to be noted that when $e_i = 0$, the minimum value of μ_i is a positive constant α :

$$\alpha = \left[1 + \frac{M_{\max}}{F_d} \right]^{-F_e} \quad (4)$$

so μ_i lies in the interval $[\alpha, 1]$.

m_d is called crossover point when it satisfied $\mu_i = G(e_d) = 0.5$. In this work, we let the average value \bar{M} of the whole neighbor list as the crossover point, $e_d = \bar{M} = \frac{1}{n} \sum_{i=1}^n \frac{e_i}{dtoBS_i}$. Then the constant

F_d can be calculated as

$$F_d = \frac{M_{\max} - \bar{M}}{\sqrt{2} - 1} \quad (5)$$

Step3: The fuzzy set U can be modified to enhance contrast of the membership through the non-linear transform T :

$$\mu'_i = T^{(p)}(\mu_i) = T(T^{(p-1)}(\mu_i)) \quad (p = 1, 2, \dots) \quad (6)$$

where p is the iterative times of T . By (6), we can get the new fuzzy property plane $U' = \{\mu'_i, i = 1, \dots, n\}$.

In [11], T is a kind of S-shaped function and the fuzzy enhancement function is given by

$$T(\mu_i) = \begin{cases} 2(\mu_i)^2, & 0 \leq \mu_i \leq 0.5 \\ 1 - 2(1 - \mu_i)^2, & 0.5 \leq \mu_i \leq 1 \end{cases} \quad (7)$$

Fig. 1 shows the relationship between U and U' , from Fig. 1 we can see the membership is increased for $\mu_i > 0.5$ and decreased for $\mu_i < 0.5$. At the crossover point, the membership value is fixed at 0.5. Thus the contrast is enhanced between nodes whose ratio is larger than \bar{M} and nodes whose ratio is smaller than \bar{M} . We will use this property of the membership which is mapped from node ratio to help select CHs in next step.

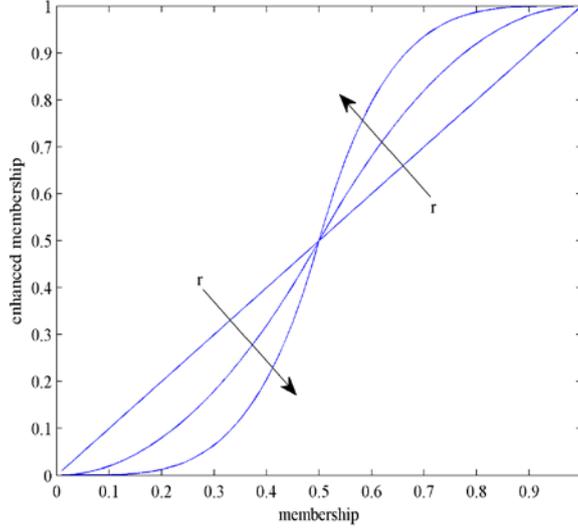


Fig.1. Relationship between U and U'

Step4: Using a converting function, node i maps the enhanced membership μ'_i into an initial time t_i as

$$t_i = \ln \frac{1+x}{\mu'_i} \quad (8)$$

where x is a random variable with uniform distribution on the interval $[0,1]$, and its role is to reduce the probability that collisions happen when two nodes may get the same.

From (8), the maximum value of t_i is easily got as

$$t_{\max} = \ln \frac{2}{\alpha} \quad (9)$$

thereafter, t_i is unitized and node i gets its waiting time t_{wt}^i

$$t_{wt}^i = T_{CH} \times \frac{t_i}{t_{\max}} \quad (10)$$

where T_{CH} is a pre-determined system parameter standing for the maximal time for CH selection.

Due to $0 < \mu'_i \leq 1$, From (8), we can get $t_i > 0$ and the first-order derivate of the function $t_i(\mu'_i)$ is $\frac{dt_i}{d\mu'_i} = -\frac{1}{\mu'_i} < 0$, so t_i is monotonously decreasing when the membership μ'_i is increasing. That is

to say, an energy-rich node which has low communication cost to the BS will have a shorter time and is more likely to become a CH.

B. Cluster Formation and Data Transnission

In this phase, a non-CH node decides on which cluster to join. Since any non-CH node may

receive one or more ADV_CH messages, such a node attempts to associate itself with the CH with highest residual energy by sending a JOIN_REQ message.

After cluster formation, A non-CH sends its collected data to the CH node which is selected in the cluster formation phase. After receiving data from all of its member nodes, the CH performs data aggregation and then transmits a packet to the BS.

Experimental Results

A. Simulation Parameters

Table.1. Simulation Parameters

| Parameter | Value |
|------------------------------|---|
| Network field | From (0,0) to (100,100) |
| Node numbers | From 100 to 500 |
| Base station | At (50,175) |
| Initial energy | $E_{\min} = 0.1\text{J}$, $E_{\max} = 0.5\text{J}$ |
| Data packet size | 525 bytes |
| Broadcast packet size | 25 bytes |
| E_{elec} | 50nJ/bit |
| ϵ_{fs} | 10pJ/bit/m ² |
| ϵ_{mp} | 0.0013pJ/bit/m ⁴ |
| E_{DA} | 5nJ/bit/signal |
| Cluster radius | 25m |
| Threshold distance (d_0) | 75 m |
| Iterative times p | 2 |

We evaluate the performance of the proposed algorithm and compare it with LEACH, Backoff Strategy in terms of the network lifetime and data capacity. The initial energy allocation on each node follows a uniform distribution. To compare the performance of DFSC with other protocols on networks of different sensor density, we let the number of sensor nodes ranging from 100 to 500. Table 1 lists all the simulation parameters. In each experiment, we run the protocols 100 times where each time uses a random topology of sensor nodes.

B. Simulation Results

Fig. 2 shows the number of alive nodes over the simulation time when the number of nodes is 100. As shown in Fig. 2, DFSC effectively improved the network lifetime, compared to LEACH and Backoff Strategy. This is because DFSC uses the new membership metric as the parameter to select cluster heads. A node with higher residual energy and lower communication cost is selected as a CH, so the time when the first node dies is prolonged.

Fig. 3 shows the total number of data messages received at the BS over the number of rounds of activity. The plot clearly illustrates that DFSC can deliver significantly more data to be forwarded to BS than LEACH and Backoff Strategy.

Fig. 4 compares the network lifetime of the three protocols for the number of nodes ranging from 100 to 500. The figure shows that DFSC can extend the network lifetime better than LEACH and Backoff Strategy under different network scenarios.

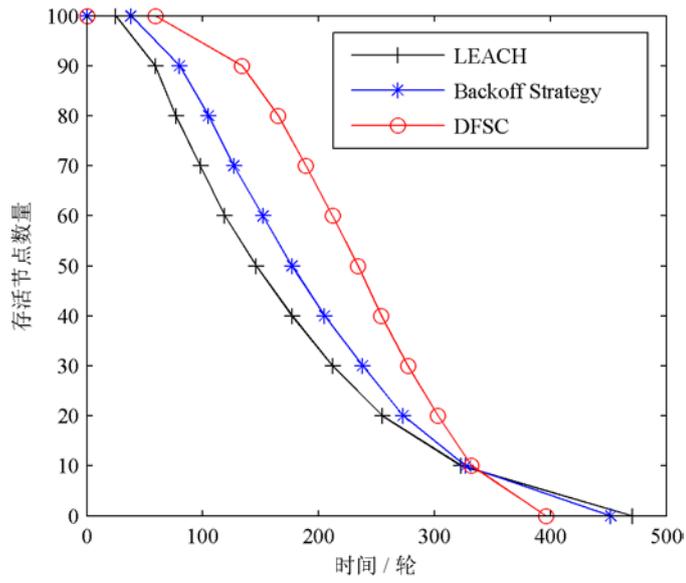


Fig.2. Number of nodes alive over time

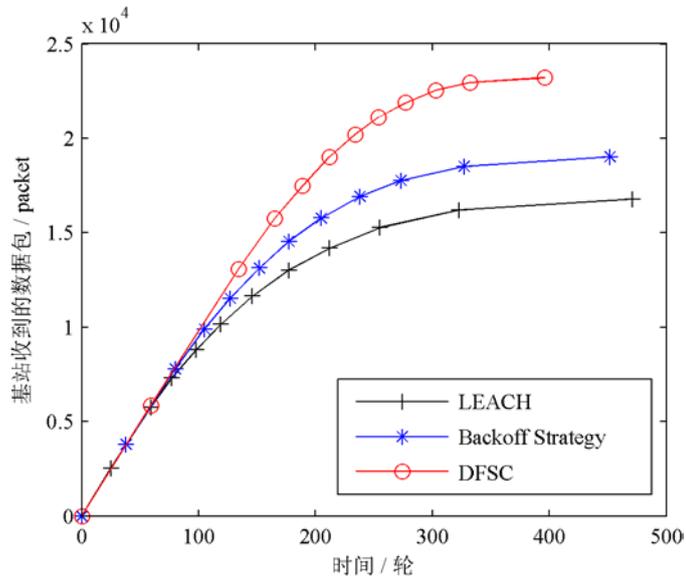


Fig.3. Number of message received in base station over time

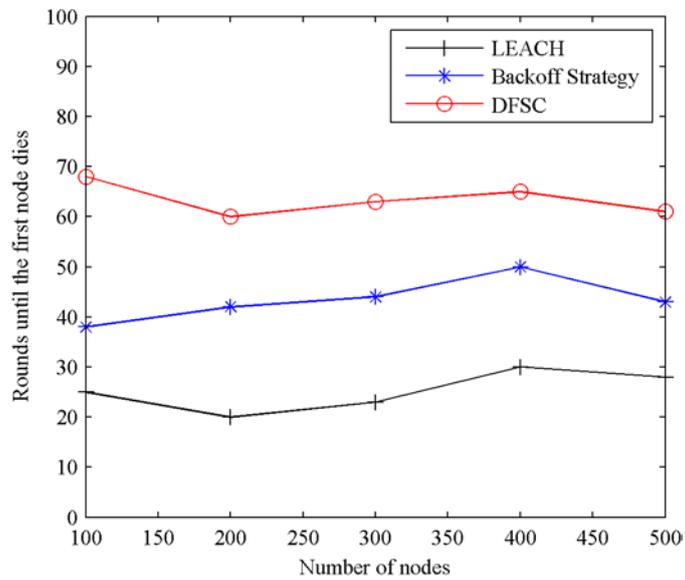


Fig.4. Network lifetime at different sensor density

Conclusion

In this paper, a novel distributed energy-efficient clustering algorithm called DFSC for multilevel energy heterogeneous wireless sensor networks is introduced. A node is selected to be a CH based on its membership which considering both node residual energy and the distant to the base station. The numerical results show that the proposed algorithm could better handle heterogeneous energy circumstances and prolong the network lifetime.

References

- [1] I. F. Akyildiz, W. Su, and Y. Sankarasubramaniam, "A survey on sensor networks," *IEEE Communications Magazine*, vol. 40, pp. 102-114, August 2002.
- [2] W. B. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Transactions on Wireless Communications*, vol. 1, pp. 660-670, October 2002.
- [3] G. Smaragdakis, I. Matta, and A. Bestavros, "SEP: A Stable Election Protocol for clustered heterogeneous wireless sensor networks," in *Proceedings of 2nd International Workshop on Sensor and Actor Network Protocols and Applications (SANPA 2004)*, pp 251-261. 2004.
- [4] L. Qing, Q. Zhu, and M. Wang, "Design of a distributed energy-efficient clustering algorithm for heterogeneous wireless sensor networks," *Computer Communications*, Vol. 29, pp. 2230-2237, March 2006.
- [5] Y. Cao and C. He, "A distributed clustering algorithm with an adaptive backoff strategy for wireless sensor networks," *IEICE Transactions on Communications*, vol. E89-B, pp. 609-613, February 2006.
- [6] O. Younis and S. Fahmy, "HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks," *IEEE Transactions on Mobile Computing*, vol. 3, pp. 366-379, October-December 2004.
- [7] C. Y. Wen, and W. A. Sethares, "Automatic decentralized clustering for wireless sensor networks," *EURASIP Journal on wireless communications and Networking*, Vol. 2005, pp. 686-697, 2005.
- [8] M. Youssef, A. Youssef, and M. Younis, "Overlapping multi-hop clustering for wireless sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 20, pp. 1844-1856, December 2009.
- [9] J. Xu, Y. He, and R. Li, "An energy-aware distributed clustering algorithm in wireless sensor networks," in *Proceedings of International Conference on Computer Science and Software Engineering (CSSE 2008)*, pp. 528-531, 2008.
- [10] J. M. Kim, S. H. Park, Y. J. Han, and T. M. Chung, CHEF: Cluster head election mechanism using fuzzy logic in wireless sensor networks, in *Proceedings of 10th International Conference on Advanced Communication Technology (ICACT'08)*, pp. 654-659, GangWo-Do, Korea, February 2008.
- [11] Pal S. K, King R. A, "Image enhancement using smoothing with fuzzy sets," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 7, pp. 494-501, 1981.