

Dynamic Entropy Based Combination Weighted Clustering Approach for High-Speed Ad hoc Network

Jianli Xie, Cuiran Li* and Hui Zhou

School of Electronics and Information Engineering, Lanzhou Jiaotong University
Corresponding Author: licr@mail.lzjtu.cn

Abstract

Weight based clustering has become the mainstream clustering algorithm in low-speed Ad hoc networks because of its excellent cluster stability. However, due to the dynamic topology changing in high-speed Ad hoc network, the cluster stability (network stability) decreased and the cluster maintenance costs increased sharply. To solve the problem, we propose a dynamic entropy based combination weighted clustering approach (DECW). First, according to the history messages of an evaluation node in the network, the upper bound and the lower bound value of each clustering index will be recorded, so the information entropy deviation of the indexes and dynamic entropy weight of each node can be obtained. After, the linear combination weights set of evaluation nodes is modeled as the second-order norm game, and the weight vector deviation is minimized as the optimization goal to get the multi-node dynamic entropy weights. In the cluster maintenance, a new Monte Carlo optimization is proposed to avoid the frequent cluster-heads (CHs) replacement induced of high node mobility of. Simulation results reveal that the proposed approach has the better adaptability in high-speed mobile environment.

Keywords: *Ad hoc network; Weighted Clustering; Combination weight; dynamic entropy; Monte Carlo optimization*

1. Introduction

Clustering is one of the most widely investigated solutions for scaling down ad-hoc networks. Increasing network capacity and reducing the routing overhead through clustering brings more efficiency and effectiveness to scalability in relation to node numbers and the necessity for high mobility. Cluster formation includes arranging network nodes into logical groups with the goal of cutting the signaling overhead required for network operation while upholding the network connectivity. The specific objective of the grouping process generally depends on network characteristics and application requirements. Nowadays, the researches of clustering algorithm of Ad Hoc networks focus on the weight based method, in which the value of each weight is calculated based on the importance of different factors (indexes). The methods of weighted mean can be divided into two categories roughly, the stable weight method, and the dynamic weight one. The former sets each decision attribute with stable weight in the fusion model. Once weights are set, they are difficult to be adjusted. A weight based distributed clustering algorithm (WCA) which can dynamically adapt itself with the ever changing topology of ad hoc networks was proposed in [1]. It has the flexibility of assigning different weights and takes into account a combined effect of the ideal degree, transmission power, mobility and battery power of the nodes. In [2], the authors introduced a new algorithm called Enhancement on Weighted Clustering Algorithm (EWCA) to improve the load balancing, and the stability in the MANET. The cluster head is selected efficiently based on these indexes, like high transmission power, transmission range, distance mobility, battery power and energy. Anyhow, stable weight settings cannot adapt to variations of node

behaviors. In latter one, attribute weights are dynamically set according to behaviors of network entities, but there exist two problems: how to decide weights of different kinds of attributes objectively and how to adjust attribute weights dynamically according to variations of node behaviors.

According to information entropy theories, entropy describes the uncertainty extent of information included in an attribute. Its gain indicates the decrease extent of the uncertainty using the attribute to divide a sample set. Attribute that has high entropy gain is thought to be informative. So it should possess high weights in decision models. Recently, some Entropy-based Weighted Clustering Algorithms are proposed, considering the entropy a better indicator of the stability and mobility of the ad hoc network[3,4], and can better reflect the objectivity of the weights. However, in the above algorithms, the weights of indexes are pre-determined without considering the time varying network topology.

In this paper, we propose an improved weight based clustering algorithm, namely dynamic entropy based combination weighted clustering approach (DECW). In the approach, according to the history messages of an evaluation node in the network, the upper bound and the lower bound value of each clustering index will be recorded to track the time varying network topology, so the information entropy deviation of each factor and dynamic entropy weight of each node can be obtained. Then, the linear combination weights set of evaluation nodes is modeled as the second-order norm game, and the weight vector deviation is minimized as the optimization goal to get the multi-node dynamic entropy weights. In the cluster maintenance, a new Monte Carlo optimization is proposed to avoid the frequent CHs replacement caused by the high mobility of nodes.

The rest of the paper is organized as follows. The details of the proposed DECW approach are presented and a new Monte Carlo optimization for cluster maintenance is given in Section 2. In Section 3, the approach's performances are evaluated and discussed. Finally, some conclusions are drawn in Section 4.

2. Dynamic Entropy Based Combination Weighted Clustering

This section describes the proposed DECW approach, which includes three major parts: determining the clustering indexes, CH selection by calculating the dynamic entropy based combination weight, and cluster maintenance using a new Monte Carlo method.

2.1 Clustering Indexes

Four indexes are taken into consideration as clustering indexes for high-speed Ad hoc networks.

- 1) Difference between the degree of each node and the ideal degree.

The neighbors of node i (i.e., nodes within its transmission range) is defined as the degree of node i , d_i

$$d_i = |N(i)| = \sum_{i \in V, i \neq i} \{ \text{dist}(i, i') < TX_{range} \} . \quad (1)$$

where V is the set of nodes, and TX_{range} is the transmission range of node i . The degree difference from the ideal degree for each node is computed as

$$\Delta_i = |d_i - d_{ideal}| . \quad (2)$$

where d_{ideal} is the number of nodes that a cluster-head (CH) can handle ideally, which is a pre-defined threshold to ensure that CHs are not over-loaded and the efficiency of the network is maintained at the expected level.

- 2) Distance of each node from the ideal CH

In view of energy consumption, if CH is located on the border of a cluster, the energy consumption is 2.4 times as well as that of selecting the nodes in the center [5]. Accordingly, the ideal CH should be selected near the center of their cluster as well as possible for energy saving. In VANET, nodes are distributed randomly, and can not guarantee that the CH is exactly located the center of a cluster.

In the n -dimensional Euclidean space, the distance of each node from the ideal CH in a cluster, can be expressed by

$$S_i = \sqrt{\sum_{j=1}^n (i_j - i_j^*)^2} \quad (3)$$

where $i = \{i_1, i_2, \dots, i_n\}$ and $i^* = \{i_1^*, i_2^*, \dots, i_n^*\}$ are n -dimensional coordinates of node i and the ideal CH i^* . For the 2-dimensional Euclidean space, (3) can be rewritten as

$$S_i = \sqrt{(i_1 - i_1^*)^2 + (i_2 - i_2^*)^2} \quad (4)$$

3) Clustering stability

In order to avoid frequent CH changes, it is desirable to elect a CH that does not move very quickly. Our clustering procedure achieves the stability by considering the relative mobility of nodes.

The average distance for every node i from its neighbor u till current time T , is calculate as

$$\bar{d}_{iu} = \frac{1}{T} \sum_{t=1}^T d_{iu}^t \quad (5)$$

where d_{iu}^t is the distance between node i and u at time t .

Let LS_{iu} represent the link stability between i and u , expressed by

$$LS_{iu} = \frac{1}{T} \sum_{t=1}^T (d_{iu}^t - \bar{d}_{iu})^2 \quad (6)$$

The average of the link stability for every node i with all its neighbors, is given by [6]

$$LStab_i = E(LS_{iu} | u \in N(i)) = \frac{1}{d_i} \sum_{u \in N(i)} LS_{iu} \quad (7)$$

Then, the cluster stability can be obtained as

$$CStab_i = \frac{1}{d_i} \sum_{u \in N(i)} E(LS_{iu} - LStab_i)^2 \quad (8)$$

From (8), we can see that the $CStab_i$ is somewhat like that of variance, which reflects the relative mobility of the nodes.

4) Node energy consumption

To develop more energy-efficient green communication systems, the energy consumption of each node should be an important index for CH election. It helps in efficient MAC functions and load balancing because it is always desirable to elect a node as CH with more remaining energy (i.e., less energy consumption).

The energy consumption of node i can be expressed by

$$E_i = e \cdot M_i \quad (9)$$

where e is the energy consumption for delivering a packet, and M_i is the number of packets delivered by node i .

2.2 CH Election

For the above four clustering indexes, denoted as A, B, C, and D respectively, each node has m ($m \geq 1$) corresponding message record, which are used to record the possible values of each index.

It is assumed that there are N evaluation nodes in Ad hoc network. For node i ($i=1,2,\dots,N$), the range of index D ($A/B/C$) values can be denoted as $[V_{iD_min}, V_{iD_max}]$ ($0 \leq V_{iD_min} \leq V_{iD_max}$). By the principle of objective weighting method, the larger deviation between some index values, the more important it is in algorithm selection or sorting, so, the weight of the index should be allocated a larger value. Otherwise, it has the smaller weight. For node i ($i=1,2,\dots,N$), the deviation of clustering indexes is mainly determined by the upper bound and lower bound value.

For node i ($i=1,2,\dots,N$), the information entropy of clustering index D is denoted as E_{iD} , which is defined as

$$E_{iD} = -(\ln 2)^{-1} (P_{iD_min} \ln P_{iD_min} + P_{iD_max} \ln P_{iD_max}) \quad (10)$$

where P_{iD_min} and P_{iD_max} are expressed respectively, as follows

$$P_{iD_min} = \frac{V_{iD_min}}{V_{iD_min} + V_{iD_max}} \quad (11)$$

$$P_{iD_max} = \frac{V_{iD_max}}{V_{iD_min} + V_{iD_max}} \quad (12)$$

Then, the dynamic weight of index D is

$$w_{iD} = \frac{1 - E_{iD}}{4 - (E_{iA} + E_{iB} + E_{iC} + E_{iD})} \quad (13)$$

where E_{iA} , E_{iB} , and E_{iC} represent the information entropy of clustering index A, B, and C, respectively, whose expression is similar to (10). Similarly, the dynamic entropy weight w_{iA} , w_{iB} , and w_{iC} could be calculated, which satisfy the following constraints

$$\begin{cases} 0 \leq w_{iD} (w_{iA}, w_{iB}, w_{iC}) \leq 1 \\ w_{iA} + w_{iB} + w_{iC} + w_{iD} = 1 \end{cases} \quad (14)$$

So, the dynamic entropy weight vector for node i ($i=1,2,\dots,N$) is denoted as $w_i [w_{iA}, w_{iB}, w_{iC}, w_{iD}]$.

To improve the objectivity of multi-attribute weight assignment, the dynamic entropy weight vectors for all N nodes are linearly combined. Then, the multi-node dynamic entropy weight vector is defined as

$$w^* = \sum_{i=1}^N \alpha_i \cdot w_i \quad (15)$$

where α_i ($i=1,2,\dots,N$) is the combination coefficient. To find the most satisfactory weight vector w from all possible weight vector w^* , the N linear combination coefficients obtained from (15) is optimized. With the optimization goal of making w^* minimum deviation with each w_i , the following strategy model is established [7]

$$\min \left\| \sum_{j=1}^N \alpha_j w_j^T - w_i^T \right\|_2 \quad (i = 1, 2, \dots, N) \quad (16)$$

The strategy model in (16) is a second-order norm. Due to involving square root, using partial derivative is difficult to find the combination coefficient, α_i ($i=1,2,\dots,N$). Therefore, the model is converted to the form of second-order norm squared, which is written as

$$\begin{aligned} & \left\| \sum_{j=1}^N \alpha_j w_j^T - w_i^T \right\|_2^2 \\ & = \left(\sum_{j=1}^N \alpha_j w_j - w_i \right) \left(\sum_{j=1}^N \alpha_j w_j^T - w_i^T \right) \quad (i=1,2,\dots,N; j=1,2,\dots,N). \end{aligned} \quad (17)$$

To obtain the optimum $\alpha_i(i=1,2,\dots,N)$, the partial derivative of (17) is calculated, expressed by

$$\begin{aligned} & \frac{\partial}{\partial \alpha_i} \left[\left\| \sum_{j=1}^N \alpha_j w_j^T - w_i^T \right\|_2^2 \right] \\ & = w_i \left(\sum_{j=1}^N \alpha_j w_j^T - w_i^T \right) + \left(\sum_{j=1}^N \alpha_j w_j - w_i \right) w_i^T \\ & = 2 \sum_{j=1}^N \alpha_j w_i w_j^T - 2w_i w_i^T \quad (i=1,2,\dots,N; j=1,2,\dots,N) \end{aligned} \quad (18)$$

Namely,

$$2 \sum_{j=1}^N \alpha_j w_i w_j^T - 2w_i w_i^T = 0. \quad (19)$$

Then, we can get

$$\sum_{j=1}^N \alpha_j w_i w_j^T = w_i w_i^T. \quad (20)$$

And, (20) can be rewritten as

$$\begin{pmatrix} w_1 \cdot w_1^T & w_1 \cdot w_2^T & \dots & w_1 \cdot w_n^T \\ w_2 \cdot w_1^T & w_2 \cdot w_2^T & \dots & w_2 \cdot w_n^T \\ \vdots & \vdots & \dots & \vdots \\ w_n \cdot w_1^T & w_n \cdot w_2^T & \dots & w_n \cdot w_n^T \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} w_1 \cdot w_1^T \\ w_2 \cdot w_2^T \\ \vdots \\ w_n \cdot w_n^T \end{pmatrix}. \quad (21)$$

So, the coefficient $(\alpha_1, \alpha_2, \dots, \alpha_n)$ can be obtained. And, the multi-node dynamically entropy weight vector is expressed by

$$w = \sum_{i=1}^N \alpha_i \cdot w_i. \quad (22)$$

Combined the multi-node dynamically entropy weight, $w[w_A, w_B, w_C, w_D]$, with the subjective weight, $(S[S_A, S_B, S_C, S_D])$, then, the combination weight is obtained by

$$W(W_A, W_B, W_C, W_D) = \beta_1 S + \beta_2 w. \quad (23)$$

Where $S_A, S_B, S_C,$ and S_D is the subjective weight of index $A, B, C,$ and D , whose value should be set according to the different applications scenarios in Ad hoc network. β_1 and β_2 are the combinatorial weight coefficients, which satisfy

$$\beta_1 + \beta_2 = 1 \quad 0 \leq \beta_1 \leq 1, 0 \leq \beta_2 \leq 1. \quad (24)$$

Then, the combinatorial weight of clustering indexes for node $i(i=1,2,\dots,N)$, is expressed by

$$W_i = W_A \Delta_i + W_B S_i^* + W_C CStab_i + W_D E_i. \quad (25)$$

The node with the smallest W would be selected as CH. If more than one node has the smallest W , the node with better clustering stability is assigned as CH. All the neighbors of the chosen CH are no longer allowed to participate in the CH election procedure.

2.3 Cluster Maintenance

The node mobility often brings the changing in communication state, which results in cluster updating. In high-speed Ad hoc network, when a node moves into a cluster, and its weight is lower than that of the CH, then CH competition is inevitable. In wireless communication environments, the changing in communication state induced by the node mobility, interference and multipath fading etc., might result in the weight jitter frequently. That is, when the weight difference is small between two nodes, the phenomenon of swapping weight ordering would happen frequently, due to slight weight jitter. In this case, it is reasonable to remain the CH unchanged in a larger probability

Monte Carlo is a method obtained from the drop rule of local search algorithm^[8], in which the basic idea is to accept a relatively worse state in a certain probability. The Monte Carlo method in cluster maintenance is as follows^[9]: when CH- i encounter competition of cluster-member (CM)- v ($W_i > W_v$), the CH will be replaced in probability $(1-p)$; the accepting probability is defined as $p = \min\{1, e^{(W_i - W_v)/v_{\max}}\}$, and the CH which has the larger weight continues to act as a CH in probability p . When the variable x is less than 0, exponential function $y=e^x$ decrease monotonically. If x is infinitely small, y is close to 0. The accepting probability is defined as exponential function, which satisfies: 1) the larger weight difference between i and v , the larger replacing probability; 2) $p = 1$, if $W_i = W_v$.

In high-speed Ad hoc network, the high mobility of nodes will lead to network topology changing frequently, which has a great impact on network stability. Therefore, it is necessary to take the node mobility into consideration in cluster maintenance, and a new accepting probability is defined as

$$p' = \begin{cases} 1 - \min\{1, e^{v_{\max}(W_v - W_i)/(v_i - v_v)}\} & v_i > v_v \\ \min\{1, e^{(W_v - W_i)/v_{\max}}\} & v_i = v_v \\ \min\{1, e^{v_{\max}(W_v - W_i)/(v_v - v_i)}\} & v_i < v_v \end{cases} \quad (26)$$

where v_i and v_v are the mobile speed of node i and v , respectively, and v_{\max} is the maximum speed of a node. By the new Monte Carlo optimization method, we can see that when the weight difference is very small between two nodes, and the CH has the lower mobility, then CH is updated in a small probability. Thus, the frequent cluster updating induced of high mobility might be avoided and the cluster stability is then enhanced.

3. Performance Evaluations

The simulation parameters are listed in Table 1.

Table 1. Main Simulation Parameters

Parameters	Value
Radio propagation model	Shadowing model
Network size	500m×500m
Transmission range	$r = 50m, 100m, 150m, 200m, 250m, 300m$
Node mobile speed	$[v_{\min}, v_{\max}] = [10m/s, 100m/s]$
Subjective weight	$S[0.2, 0.3, 0.3, 0.2]$
MAC protocol	IEEE802.11 DCF
Initial node energy	50J
Simulation time	600s

Figure 1 shows the comparison of the average number of CHs with varying transmission range for different clustering algorithms. We observe that the average number of CHs decreases with the increasing transmission range. This is due to the fact that a node with a larger transmission range will cover a larger coverage area, so the number of nodes that the cluster could accommodate increases and the average number of CHs is decreased. However, for small transmission range, most nodes remain out of each other's transmission range, the probability of alone cluster increases, thus the average number of CHs is increased. Also, the results show that the proposed DECW approach produces fewer clusters than CBMD algorithm [10].

Figure 2 depicts the variation of the average number of CHs with respect to the node speed. With the increasing speed, the average number of CHs changes frequently in a low variation. This is because the number of nodes and the node mobile speed are set some value in advance for clustering process. Furthermore, it is observed that, the more number of nodes, the more average number of CHs, and DECW produces fewer clusters than CBMD algorithm.

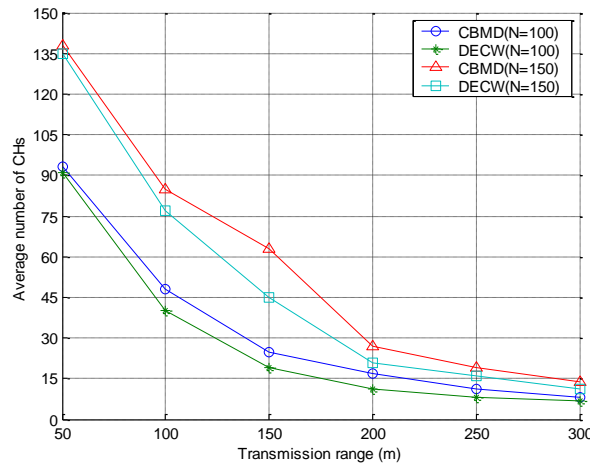


Figure 1. The Comparison of the Average Number of CHs ($v_{max}=100m/s$)

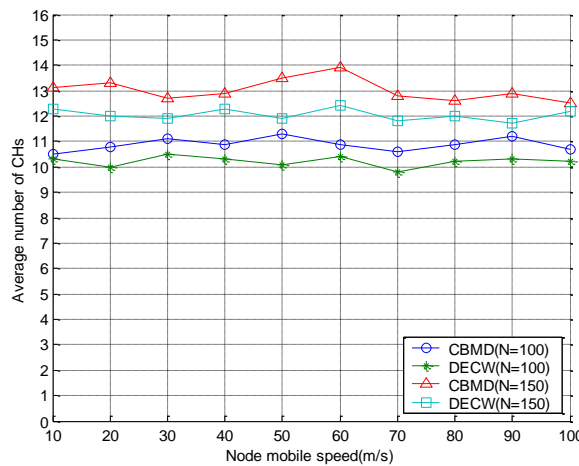


Figure 2. Average Number of CHs vs. Node Speed ($r=200m$)

Figure 3 and Figure 4 depict the variation of number of inter-cluster transferring and the numbers of CH updating with respect to the transmission range, respectively. We can see that for small transmission range, the probability of alone cluster increases, results the fewer number of CHs. With the increasing transmission range, the number of nodes that the cluster could accommodate increases, and the probability of a node moves out of the transmission range of a cluster becomes larger. So, the number of inter-cluster transferring increases, also the numbers of CH updating. When the transmission range increases to a threshold value (250m), with further increasing transmission range, the probability of a node moves out of current cluster decreases, the number of inter-cluster transferring and the numbers of CH updating decrease.

Figure 5 and Figure 6 depict the variation of the number of inter-cluster transferring and the numbers of CH updating with respect to node speed, respectively. It is noticed that number of inter-cluster transferring and the numbers of CH updating increase with the higher node mobile speed. This is because the nodes with higher speed quit the current cluster and into another one more rapidly.

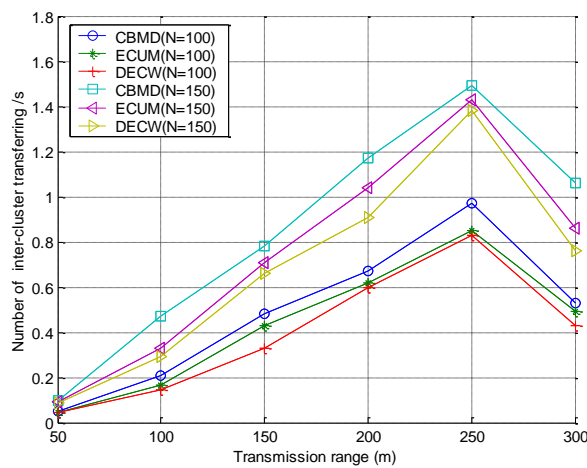


Figure 3. Number of Inter-cluster Transferring vs. Transmission Range (vmax=100m/s)

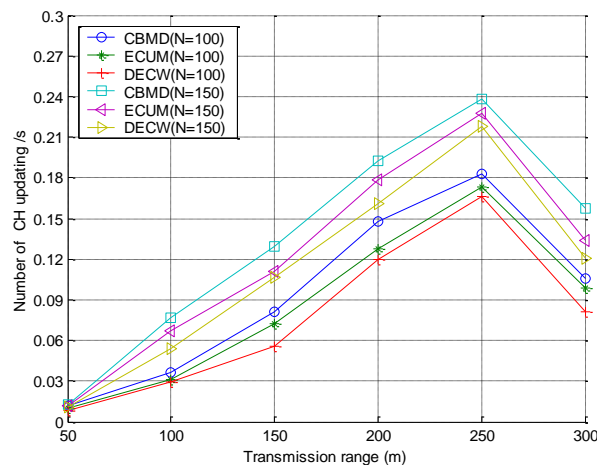


Figure 4. Numbers of CH Updating vs. Transmission Range (vmax =100m/s)

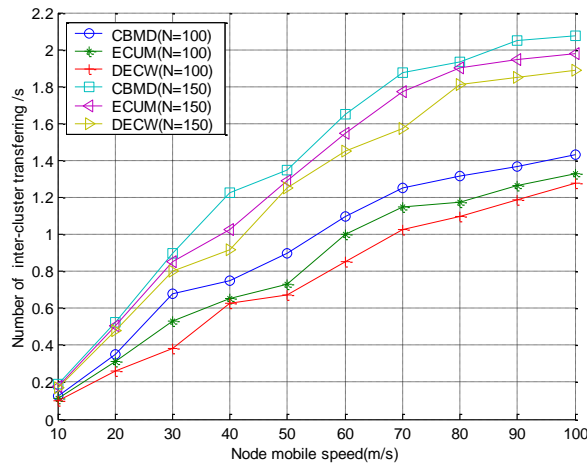


Figure 5. Number of Inter-cluster Transferring vs. Node Speed ($r=200m$)

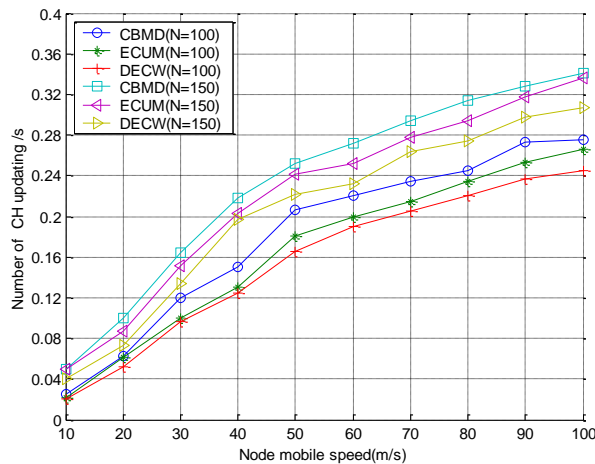


Figure 6. Numbers of CH Updating vs. Node Speed ($r=200m$)

As we observed from Figure 3-6, the cluster stability of DECW approach is better than that of CBMD and ECUM clustering algorithms [10, 11] (less number of inter-cluster transferring and the numbers of CH updating). It is the reason that DECW approach is dynamic entropy combination method, which can track the timely varying network topology more efficiently. In DECW approach, it is not only consider the difference of each clustering index for each node at some time t , but also consider the history differences of each index for each node, which makes the weight calculation more closer to the real network distribution. Besides, in cluster maintenance, we take the node mobility as an important factor to decide whether the CH is updated or not, and a new accepting probability is defined, which avoid the frequent CH updating induced by high node mobility.

4. Conclusions

To overcome the defect that network stability is poor, when the nodes move at high speed in Ad hoc network. In this paper, we presented a dynamic entropy based combination weighted clustering approach (DECW). Simulation results show that the DECW approach has the better performances than that of CBMD and ECUM clustering

algorithms in terms of the average number of CHs, the average cluster head, the number of inter-cluster transferring, and the numbers of CH updating. Also, it shows that DECW approach has the better applicability in large-scale and high-speed network environments.

Acknowledgements

This work was partially supported by Nation Science Foundation of China (61261014).

References

- [1] M. Chatterjee, S. K. Das and D. Turgut, "WCA: A Weighted Clustering Algorithm for Mobile Ad Hoc Networks", *Journal of Clustering Computing*, vol.5, no.2, (2002), pp.193-204.
- [2] F. Djemili Tolba, D. Magoni and P. Lorenz, "A stable clustering algorithm for highly mobile Ad Hoc networks", *Proceedings of Second International Conference on Systems and Networks Communications*, (2007); Cap Esterel, France.
- [3] Y. X. Wang and F. S. Bao, "An entropy-based weighted clustering algorithm and its optimization for ad hoc networks", *Proceedings of Third IEEE International Conference on Wireless and Mobile Computing, Networking and Communications*, (2007); New York, USA.
- [4] C. Gui, B. Sun, B. Yan and C. Xia, "An Entropy-Based Multipath Routing Algorithm in Mobile Ad Hoc Networks", *Proceedings of IEEE 4th International Conference on Wireless Communications, Networking and Mobile Computing*, (2008); Dalian, China.
- [5] J. Yang, D. Y. Zhang, Y. Y. Zhang and Y. WANG, "Cluster-Based Data Aggregation and Transmission Protocol for Wireless Sensor Networks", *Journal of Software*, vol.21, no.5, (2010), pp.1127-1137.
- [6] G. X. Jiang and Z. Y. Yang, "A Distributed Clustering Algorithm Based on δ -Cluster Stability for Mobile Ad hoc Networks", *Proceedings of IEEE 4th International Conference on Wireless Communications, Networking and Mobile Computing*, (2008); Dalian, China.
- [7] C. Daocai, M. Tao and L. Song, "Application of extension assessment method based on game theory to evaluate the running condition of irrigation areas", *Transactions of the Chinese Society of Agricultural Engineering*, vol.24, no.8, (2008), pp. 36-39.
- [8] J. Mei, D. Chen, J. Gao, Y. Gao and L. Yang, "Range-Free Monte Carlo Localization for Mobile Wireless Sensor Networks", *Proceedings of International Conference on Computer Science and Service System*, (2012); Nanjing, China.
- [9] L. Lin, X. Q. Xiao, M. Xu, and L. Q. Wei, "SWBCA: Stability-oriented weight-based clustering algorithm for VANETs", *Journal of Computer Applications*, vol. 30, no.7, (2010), pp.1712-1718.
- [10] A. Hussein, S. Yousef, S. Al-Khayatt and O. S. Arabeyyat, "An efficient weighted distributed clustering algorithm for mobile ad hoc networks", *Proceedings of International Conference on Computer Engineering and Systems*, (2010); Las Vegas, USA.
- [11] D. S. Chen, A. Babaei and P. Agrawal, "An efficient cluster update model in MANET", *Proceedings of IEEE 44th Southeastern Symposium on System Theory*, (2012); Florida, USA.

Authors

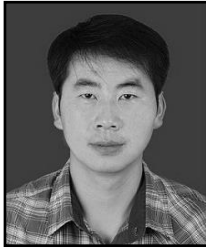


Jianli Xie received a BSc. Degree from Sichuan University in 1993, and a Master degree and Dr.degree from the School of Electronics and Information Engineering, Lanzhou Jiaotong University in 1999 and 2014, respectively. From 1999-2007, he was a senior hardware engineer in Beijing Great Dragon company. Now he is working in Lanzhou Jiaotong University as a associate professor. He has published over 10 scientific papers in his research area till now. His research interests include wireless communication network and cognitive radio techniques.



Cuiran Li received a BSc. and Master degree from the Department of Communication Engineering, Lanzhou Jiaotong University, in 1996 and 1999, respectively, and a Dr. degree from the School of Electronics and Information Engineering, Beijing Jiaotong University in 2003. She is now working in Lanzhou Jiaotong University as a professor and an IEEE member. She has published a book and over 40 scientific papers in her research area till now. Her

research interests include cognitive radio techniques, mobile ad hoc network and GSM-R/LTE-R systems.



Hui Zhou received a BSc. degree from the school of computer and communication engineering, Lanzhou University of Technology in 2011, and a Master degree from the School of Electronics and Information Engineering, Lanzhou Jiaotong University in 2014.

