Autonomous Clustering Scheme for Removing the Effects of Heterogeneous Node Degrees in Ad Hoc Networks

Ryosuke Morita
Graduate School of System Design
Tokyo Metropolitan University
6–6, Asahigaoka, Hino-shi
Tokyo, 191-0065, Japan
Email: morita-ryosuke@sd.tmu.ac.jp

Chisa Takano
Graduate School of Information Sciences
Hiroshima City University
3-4-1, Ozuka-Higashi, Asa-Minami-ku
Hiroshima, 731-3194, Japan
Email: takano@hiroshia-cu.ac.jp

Masaki Aida
Graduate School of System Design
Tokyo Metropolitan University
6–6, Asahigaoka, Hino-shi
Tokyo, 191-0065, Japan
Email: aida@tmu.ac.jp

***Abstract*—In large scale mobile ad hoc networks (MANETs), it is effective to reduce the load of routing by introducing hierarchical routing, and it is conducted by clustering of nodes. From structural constraint, clustering of nodes in MANETs should be formed by autonomous decentralized manner. A clustering mechanism based on the diffusion equation is a typical autonomous clustering in MANETs, and gives appropriate clustering if all the node degrees are uniform. However, node degrees in MANETs are heterogeneous in general, the fact causes the difference in the strength of diffusion effect. This fact causes that the position of cluster head tends to be around the boundary of networks and degrade battery efficiency of nodes. In this paper, by introducing an asymmetric diffusion depending on node degree, we propose a new clustering method independent of heterogeneity of node degrees. In addition, we show that the proposed method has efficient characteristics for battery consumption.**

1. Introduction

MANETs are composed of wireless terminals (nodes) that serve multihop communication without communication infrastructures [1]. If two wireless nodes are in radio transmission range, they can communicate directly, but if they are out of the range, they communicate via relaying the communication by intermediate nodes. This is called multihop communication. In multihop communication, routing mechanism is one of most important technical issues and the most representative and simple routing mechanisms are based on flooding of route-finding packets. Typical examples of flooding based routing are Ad hoc On-Demand Distance Vector (AODV) [2] and Optimized Link State Routing (OLSR) [3]. These methods broadcast control packets into the whole network to build communication path. Thus, in case of large scale networks with a lot of nodes, control packets increase and cause unnecessary load on networks. That is, these methods are not scalable. Effective solution to keep scalability is clustering that divides networks into some groups, and actual clustering methods are proposed [4], [5]. Since a range of broadcasted route-finding packets is limited within cluster, the load of route-finding packets is reduced and scalability is maintained. A clustering mechanism based on the diffusion equation is a typical autonomous and self-configuring clustering in MANETs. In particular, a simple clustering method using combination of diffusion and back-diffusion has been proposed [6]. This method is able to form clusters with relatively few parameters, and its formation speed is fast. In addition, cluster structure can be flexibly designed. For example, a node having higher residual battery power can be designed to become clusterhead (CH) if all the node degrees are uniform. However, node degrees in MANETs are heterogeneous in general, which causes the difference in the strength of diffusion effect. This fact cause that the position of CH tends to be around the boundary of networks and degrade battery efficiency of nodes.

To solve this problem, this paper proposes a new method that can remove the effects of heterogeneous node degrees in MANETs by introducing asymmetric diffusion to clustering. This method can select CH flexibly; for example, higher degree node and/or higher residual battery power node.

This paper consists of the following sections. Section 2 describes the conventional autonomous clustering method based on diffusion and back-diffusion. In Section 3, by introducing asymmetric diffusion, we propose a new autonomous clustering method that removes the effects of heterogeneous node degrees in MANETs, In addition, we propose design method to select CH by considering terminal's residual battery power. In Section 4, the comparison of the proposed method and the conventional method is investigated through simulation. Finally, we conclude this paper in Section 5.

2. Preliminary

In this section, we outline the conventional autonomous clustering method based on diffusion and back-diffusion [6]. In addition, we explain the problem that the position of CH tends to be selected in the area of low node density.

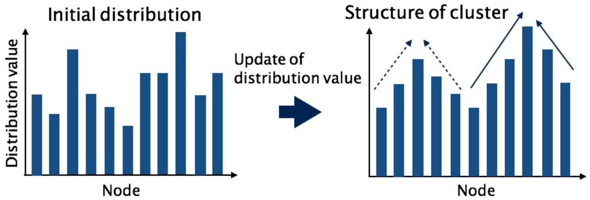

Figure 1: Outline of autonomous decentralized clustering.

2.1. Conventional Autonomous Clustering Method Based on Diffusion and Back-Diffusion

The basic procedure of the conventional autonomous clustering method is summarized as follows [6].

1) Each node has the value of state and we give an initial value of the state on the basis of network environment (e.g. the residual battery power of each node). Let the initial values of all the nodes be referred to as the initial distribution in the network.
2) Each node updates its distribution value autonomously by exchanging value with the adjacent nodes.
3) Repeating to choose the node with the largest value among the adjacent nodes, it reaches a node having a local maximum value. The node having a local maximum value is CH, and the nodes that reach the CH node by repeating the above procedure belong to the same cluster.

Since the update of the value of distribution should be conducted autonomously, the first and the second spatial derivatives (drift motion and diffusion) of the distribution governs the operation rule for update of the distribution. We design the drift motion so as to concentrate on the distribution value to the CH by using back-diffusion. Then, clusters are formed autonomously by balancing the effect of diffusion and the drift. We describe the operating rules of nodes using position x and time t of the Euclidean space in order to simplify the description. The operating rules of nodes $\boldsymbol{J}(\boldsymbol{x},t)$ of the distribution value $q(x,t)$ at time t and position x is expressed as follow:

$$\boldsymbol{J}(\boldsymbol{x},t) = -c\,\boldsymbol{f}(\boldsymbol{x},t)\,q(\boldsymbol{x},t) - c\,\sigma^2\,\nabla q(\boldsymbol{x},t), \quad (1)$$

$$\boldsymbol{f}(\boldsymbol{x},t) = -\nabla\Phi(\boldsymbol{x},t), \quad (2)$$

$$\Phi(\boldsymbol{x},t+\Delta t) = -(q(\boldsymbol{x},t) - \rho\,\nabla^2 q(\boldsymbol{x},t)\,\Delta t). \quad (3)$$

The first and the second terms on the r.h.s. of (1) mean effects of the drift motion and the diffusion, respectively. $\Phi(\boldsymbol{x},t)$ in (2) denotes a potential function of drift motion, and the strength of drift motion $\boldsymbol{f}(\boldsymbol{x},t)$ is determined by the potential. To emphasize the peak of the distribution by drift motion, the temporal evolution of the potential function is applied the back-diffusion to adjust strength of the drift in (3). The back-diffusion indicates an effect like time reversal of diffusion phenomenon. Parameters σ^2 and ρ denote the strength of diffusion and back-diffusion, respectively, and c governs the speed of temporal evolution of the distribution. They are positive constants. Since these equations do not contain third-order or higher differentiation with respect to spatial variable, the operation of nodes can be decided by using a local information about the node itself and its adjacent nodes. Therefore, the operation rules of each node is possible to apply to any network topology.

The conventional clustering method mentioned above tends to select a CH as the node having high initial value [6]. This characteristic is preferred in the following situation. If we set the initial distribution as terminal's residual battery power, the conventional clustering method selects CH that have much residual battery power. Since CH dedicates management tasks for cluster members, longer life time of CH is important for the lifetime of MANETs.

2.2. Technical Issue in the Conventional Clustering

The conventional clustering method tends to select CH in the area of the border of networks. In other words, the conventional clustering method tends to select CH of low node degree. This is because the strength of diffusion depends on the node degree, and the strength of diffusion of low degree node is relatively weak. So, low degree nodes (in low density area) tend to have large value of the distribution. In this subsection, we show details about the cause of this characteristic.

First, let a distribution value of node i ($i = 1, \ldots, n$) at time t be $q_i(t)$. The diffusion in networks is caused by this transfer of the distribution value from node i to adjacent node j. The transfer rate $J^{[i\to j]}(t)$ is written as

$$J^{[i\to j]}(t) = \kappa\,(q_i(t) - q_j(t)), \quad (4)$$

where κ (> 0) is a diffusion coefficient. When $J^{[i\to j]}(t) < 0$, it expresses transfer rate of reverse direction. In addition, this equation is an antisymmetric regarding to replace i and j, that is $J^{[i\to j]}(t) = -J^{[j\to i]}(t)$. Thus, we are possible to regard the diffusion phenomenon as transfer of distribution value between adjacent nodes. That is, the quantity of the value transferred from a node is equal to the quantity of the value receiving at the corresponding node. The diffusion on networks transfer an amount proportional to the difference of a pair of adjacent nodes, and the amount is independent of their node degrees. In other words, a node having high node degree decreases its value of the distribution, at higher rate. Therefore, the effect of diffusion is relatively weak at low degree node (at low node density area) if node degree is heterogeneous. For that reason, low degree node has weaker effect of diffusion, and these nodes tend to have local maximum value of the distribution. As a result, nodes in low density area tend to become CHs, because nodes having local maximum of the distribution are selected as CHs.

3. Proposal of Flexible Clustering Method by Using Asymmetric Diffusion

In this section, we propose a new flexible clustering method by introducing an asymmetric diffusion which depends on the node degree or residual battery power of node.

3.1. Proposal of Flexible Clustering Method by Using Asymmetric Diffusion with Respect to Node Degree

Let the distribution value of node in general network with n nodes be $\boldsymbol{q}(t) := {}^t(q_1(t), \ldots, q_n(t))$ in vector representation. Diffusion phenomena is generated by the transfer rate (4) and the diffusion equations is expressed as

$$\frac{\mathrm{d}\boldsymbol{q}(t)}{\mathrm{d}t} = -\kappa\, L\, \boldsymbol{q}(t), \tag{5}$$

where $L := D - A$ is the Laplacian matrix, and D is the degree matrix. Also, the adjacency matrix A is defined as follow;

$$A_{ij} = \begin{cases} 1 & \text{link between } i - j \text{ exists}, \\ 0 & \text{otherwise}. \end{cases} \tag{6}$$

Since the transfer rate (4) for each link is independent of the node degree, the speed of diffusion at each node depends on its node degree. Let us introduce the normalized Laplacian matrix $N := D^{-1/2}\, L\, D^{-1/2}$, and we replace L in (5) with N as

$$\frac{\mathrm{d}\boldsymbol{q}(t)}{\mathrm{d}t} = -\kappa\, N\, \boldsymbol{q}(t). \tag{7}$$

where N is a symmetric matrix as with L. Let $J_{i,j}(t)$ be the change rate of distribution value $q_i(t)$ at node i caused by the effect from an adjacent node j, and it is written as

$$J_{i,j}(t) = \kappa\left(\frac{q_i(t)}{d_i} - \frac{q_j(t)}{\sqrt{d_i d_j}}\right), \tag{8}$$

where d_i is degree of node i. This equation is not anti-symmetric with respect to replacement of i and j, that is $J_{i,j}(t) \neq -J_{j,i}(t)$. We call the phenomenon described by the temporal evolution equation (7), as an asymmetric diffusion with respect to node degrees. Thus, we cannot regard that the asymmetric diffusion is a transfer phenomena of distribution value between adjacent nodes. That is, each node changes its distribution value from the effect of asymmetric interaction between nodes [7]. We can expect that the effects of heterogeneous node degree in MANETs is removed by using the change rate. That is because this change rate can cancel of the difference of change of distribution value with respect to node degree.

We propose a new clustering mechanism by replacing the diffusion in the conventional clustering with asymmetric diffusion based on node degree. Actually, the second term on the r.h.s. of (1) is (4) in the conventional method, but is (8) in the proposed method.

3.2. Proposal of Flexible Clustering Method by Using Asymmetric Diffusion with Respect to Residual Battery Power of Node

It is necessary for the clustering mechanism to consider both CH's node degree and terminal's residual battery power. In order to realize the flexibility in clustering method, we should be able to extend the temporal evolution equation by various property of node. Let the property of node i be m_i (> 0), and $M := \mathrm{diag}(m_1, \ldots, m_n)$. In addition, we define the scaled laplacian matrix $S := M^{-1/2}\, L\, M^{-1/2}$, and we replace L in (5) with S as

$$\frac{\mathrm{d}\boldsymbol{q}(t)}{\mathrm{d}t} = -\kappa\, S\, \boldsymbol{q}(t), \tag{9}$$

where S is symmetric matrix as with L and N. The change rate $J_{i,j}(t)$ of distribution value based on (9) is written as follows:

$$J_{i,j}(t) = \kappa\left(\frac{q_i(t)}{m_i} - \frac{q_j(t)}{\sqrt{m_i m_j}}\right). \tag{10}$$

We also propose a new flexible clustering mechanism by replacing the diffusion in the conventional clustering with asymmetric diffusion based on a certain node characteristic m_i. Actually, the second term on the r.h.s. of (1) is (4) in the conventional method, but is (10) in the proposed method.

By choosing m_i as

$$m_i = {d_i}^{\alpha}\, q_i(0)^{1-\alpha}, \tag{11}$$

using node degree d_i and the residual battery power (initial value of distribution) of the node, the temporal evolution of $q_i(t)$ can be controlled by node degree and the residual battery power of the node. By adjusting parameter α in $0 \leq \alpha \leq 1$, we can adjust a balance of node degree and the residual battery power of node in clustering process. When $\alpha = 1.0$, (10) is the same with (8).

4. Evaluations of the Proposed Clustering Method

In this section, we show the comparison of the proposed method with the conventional method through simulation.

4.1. Simulation Conditions and Preliminary Evaluations

In this simulations, we use the unit disk graph (UDG) as network models. In UDG, two nodes are connected by a link if and only if the distance between them is less than the predefined value [8]. Let us draw a circle around a node such that the radius of the circle is a half of the predefined threshold of the distance as Figure 2. Two nodes have link if and only if the corresponding two circles are overlapped. This model is suitable for simulation of MANETs. This is because a link means a wireless link and the distance threshold corresponds radio range of nodes. Heterogeneity of node degree is determined by the spatial distribution of node positions.

In our evaluation, we use the network model with 500 nodes in a square plane of $[0, 1] \times [0, 1]$, and set the radio range to 0.07. In addition, we give the value of initial distribution to each node by uniform random numbers of interval $[1, 2]$. The parameters used in the conventional and the proposed methods are described in TABLE 1. To

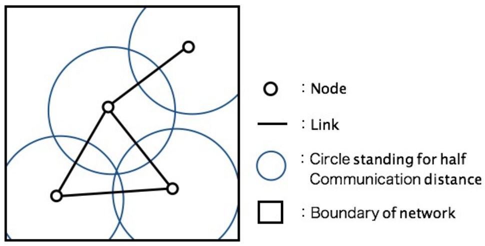

Figure 2: Unit Disk Graph.

TABLE 1: Parameters in simulation models.

Parameters	c	σ^2	ρ	α
Proposal method	0.001	80.0	0.12	—
Exiting method	0.001	40.0	0.12	0.0∼1.0

balance the convergence speed of the number of the formed clusters in both methods, the parameter σ^2 of our proposed method is set to be stronger than that of the conventional method.

Since we would like to compare different clustering methods under the condition of the same number of clusters, we show that convergence process of the number of clusters as a preliminary evaluation. Figure 3 shows the mean number of clusters with respect to the iterations of updates of distribution value, for different α. In the conventional method, the mean number of clusters gradually decrease. This is because that the diffusion acts stronger than the drift. However, in our proposed method, the mean number of clusters converges to a constant value. The convergence is realized by introducing asymmetric diffusion characterized by the property of node m_i. Then, the converged distribution values is higher for the nodes having larger value of the property. Therefore, the change from the initial distribution becomes smaller if the parameter α is small, and the converged value of the mean number of clusters becomes larger. Figure 3 shows that the mean number of clusters of our proposed method converges to about 30 when $\alpha \simeq 1$. Hence, we compare temporal evolutions of the number of clusters, obtained from both the conventional and the proposed methods not less than that the number of clusters is 30.

4.2. Node Degree of Cluster Heads

We compare the degree distribution of cluster heads obtained from the proposed method and the conventional method. Figure 4 shows the degree distribution of cluster heads under the condition that the number of clusters is 50, 40 and 30. In addition, Figure. 5 shows the average node degree of cluster heads and its 95% confidence interval. From Figure. 4, it is recognized that the conventional method tend to select CH of low node degree. In contrast, our proposed method tends to select CHs of higher node degree compared with conventional method. Figure. 5 shows that

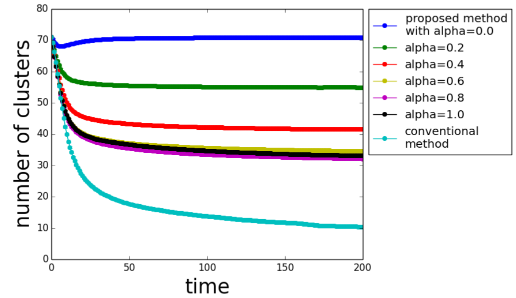

Figure 3: Dependence of mean number of cluster on time t.

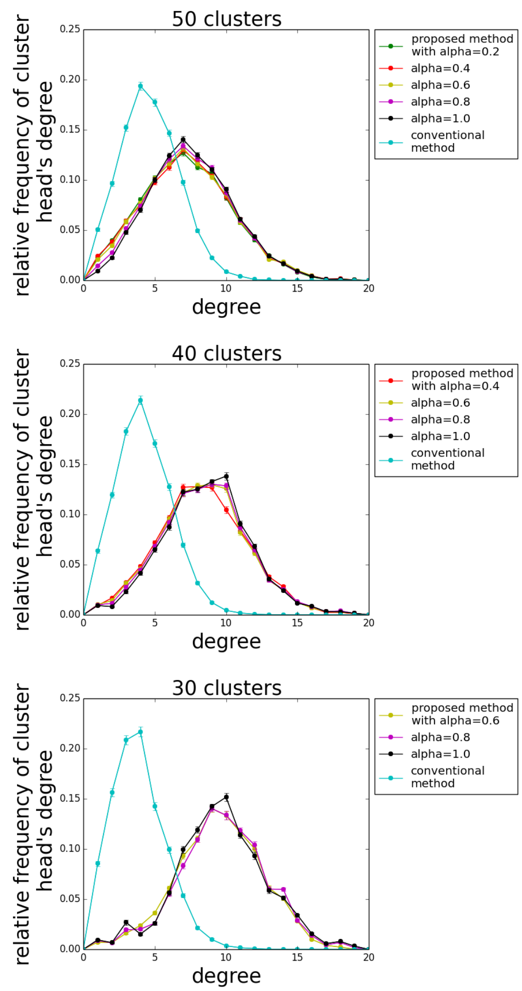

Figure 4: Degree distribution of CHs.

the average degree of CHs for the proposed method is higher than that for conventional method.

4.3. Residual Battery Power of Cluster Heads

Since CH dedicates management tasks for cluster members, longer life time of CH is important for the lifetime

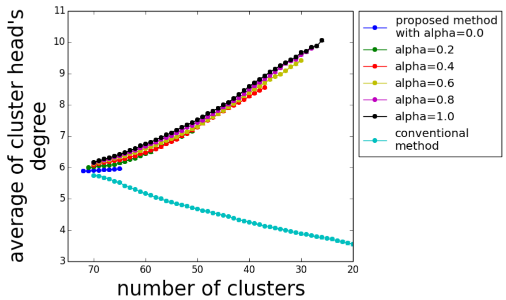

Figure 5: Average of CH's degree.

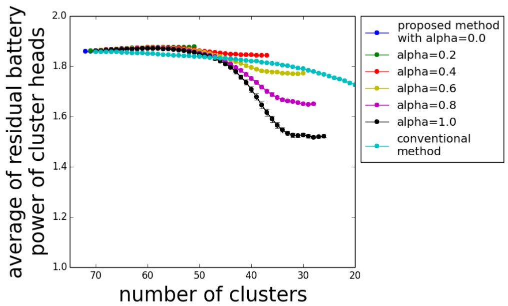

Figure 6: Average of Residual Battery Power of CHs.

of MANETs. By setting the initial distribution on nodes as terminals' residual battery power, we expect that clustering methods select CH having much residual battery power. This subsection shows the comparison of residual battery power of CHs obtained from the proposed method and the conventional method.

Figure 6 shows the average residual battery power of of CHs and its 95% confidence interval. The horizontal axis denotes the number of clusters, and the vertical axis denotes the average residual battery power pf CHs. From this figure, we can recognize that the conventional method selects CH that have much residual battery power. For the proposed method, although parameter of $\alpha = 1.0$ gives less residual power, we can obtain greater residual power if we set small value of α.

4.4. Efficiency of Intra-Cluster Routing and Node Centrality

In this subsection, we discuss hierarchical routing mechanisms after the clustering. In particular, the discussion about the efficiency of routing is based on two types of node centrality of graph.

4.4.1. Closeness Centrality and Betweenness Centrality. In hierarchical routing mechanism in MANETs, management of cluster, routing and communication are conducted throough cluster head base tree (CHBT). CH is the root of CHBT. The two types control packets are used for construction of CHBT. They are called Member Packet (MEP) and Member Ack Packet (MAP). Construction procedure of CHBT is summarized as follows.

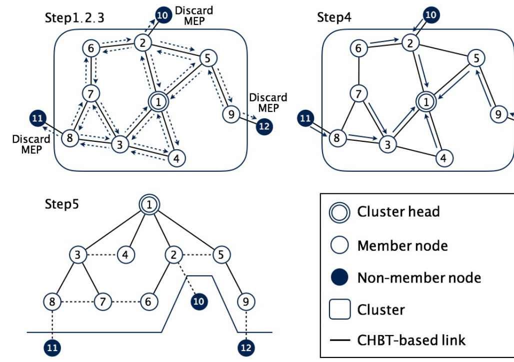

Figure 7: Image when CHBT is constructed.

1) The CH broadcasts MEP to adjacent nodes at fixed intervals.
2) The node that received MEP broadcasts MEP, until all the nodes in the cluster receive it.
3) MEP is discarded when it reaches outside of the cluster.
4) The node that received MEP sends MAP toward CH after a certain period of time.
5) CHBT complete when CH receive MAP.

MEP includes cluster ID information. Thus, by comparing own cluster ID with cluster ID of MEP, each node decides to broadcast again or to discard. Figure 7 shows the image of CHBT's construction process. Since CH has a routing table of CHBT, all the communications should be via CH. Hi-AODV [9] is one of the hierarchical routing protocol. The intra-cluster routing is conducted based on CHBT and the inter-cluster routing is based on AODV. For intra-cluster routing, the node that have high closeness centrality and high betweenness centrality is suitable as CH. Closeness centrality is the index such that the node near other nodes is high centrality. It is defined by the sum of the reciprocal of the number of shortest hops. Closeness centrality of node i in an undirected graph $G(V, E)$ that composed node set V and link set E is defined as

$$C_c(i) = \frac{N-1}{\sum_{i \in V} \mathrm{dist}(i,j)}, \tag{12}$$

where N is the number of nodes, $\mathrm{dist}(i,j)$ is the shortest distance of the number of hops between node i–j. Betweenness centrality is the index that expresses extent to which a node is present at the shortest path between other nodes. It can evaluate the mediated effect of node in communication. Betweenness centrality of node i in an undirected graph $G(V, E)$ is defined as

$$C_b(i) = \sum_{s \neq i \in V} \sum_{t \neq i \in V} \frac{\sigma_{st}(i)}{\sigma_{st}}, \tag{13}$$

where σ_{st} is the number of shortest path between node s–t, $\sigma_{st}(i)$ is the number of shortest path between node s–t including node i.

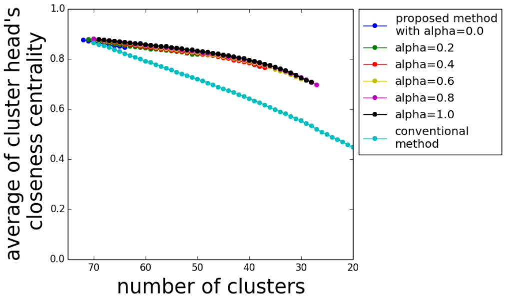

Figure 8: Closeness centrality of CHs.

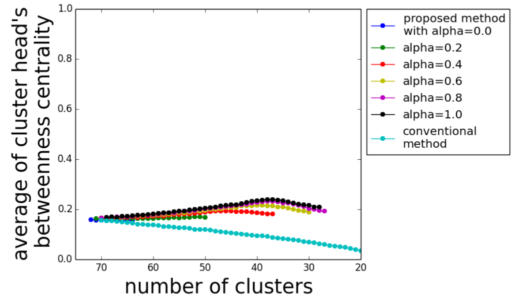

Figure 9: Betweenness centrality of CHs.

MEP is flooded from CH to nodes within cluster at the first step for constructing CHBT. In this process, the number of broadcasting of MEP can be reduced if CH has the high value of closeness centrality. This is because CH is close to other nodes. Similarly, the number of hops from any nodes to CH becomes short. In other words, the higher closeness centrality of CH gives efficient construction of CHBT. Next, for intra-cluster communication after construction of CHBT, sender node sends data to the destination via CH. Then, the nodes in cluster are able to communicate with the short number of hops if CH has the high value of betweenness centrality. Therefore, the node having the high value of betweenness centrality is suitable to CH that is relay node. This is because such CH tends to exist on the shortest path between nodes. From the above discussion, closeness centrality and betweenness centrality are important indices that evaluate the efficiency about construction of CHBT and communication within cluster.

4.4.2. Efficiency of Intra-Cluster Routing. We discuss the efficiency of intra-cluster routing by using the closeness centrality and the betweenness centrality of CH. Figures 8 and 9 show the values with 95% confidence interval of the closeness centrality and the betweenness centrality, respectively. Since the number of nodes in a cluster becomes large if the number of clusters becomes smaller, values of the closeness centrality decrease regardless clustering methods. However, Figure 8 shows the proposed method gives relatively higher value of the closeness centrarity. In addition, Figure 9 shows that the proposed method has the higher value of betweenness centrality than conventional method. Therefore, our proposed method is more effective than conventional method for intra-cluster routing.

5. Conclusion

In this paper, by introducing an asymmetric diffusion with respect to node degree, we proposed a new clustering method that is independent of heterogeneity of node degrees in MANETs. In addition, we showed that the proposed method has efficient characteristics for battery consumption. The proposed method tends to select CHs of higher node degree compared with the conventional method, and it can select CH that have greater residual battery power if we set small value of α.

In addition, the evaluation of the closeness centrality and the betweenness centrality shows that the proposed method is more effective than the conventional method for intra-cluster routing. As further work, we should consider motion of nodes and investigate the relationship between battery consumption and the efficiency of routing.

Acknowledgments

This research was supported by Grant-in-Aid for Scientific Research (B) No. 26280032 (2014–2016) from JSPS.

References

[1] C.E. Perkins, "Ad Hoc Networking," Addison Wesley, December 2000.

[2] C.E. Perkins and E.M. Royer, "Ad-hoc on-demand distance vector routing," Request for Comments 3561, 2003.

[3] T. Clausen and P. Jacquet, "Optimized link state routing protocol (OLSR)," Request for Comments 3626, 2003.

[4] E.M. Belding-Royer, "Hierarchical routing in ad hoc mobile networks," Wireless Com-munications and Mobile Computing, vol.2, no.15, pp.515–532, 2002.

[5] T. Ohta, M. Fujimoto, S. Inoue and Y. Kakuda, "Hi-TORA: Hierarchical Routing Protocol in Ad hoc Networks," 7th IEEE International Symposium on High Assurance Systems Engineering, pp.143–148, 2002.

[6] C. Takano, M. Aida, M. Murata, and M. Imase, "Proposal for Autonomous Decentralized Structure Formation Based on Local Interaction and Back-Diffusion Potential," IEICE Transactions on Communications, vol. E95-B, no. 5, pp. 1529–1538, 2012.

[7] M. Aida, C Takano, "An asymmetric node interaction model based on the scaled Laplacian matrix," Shingakugihou, NS2015-8, 2015.

[8] Brent N. Clark, Charles J. Colbourn and David S. Johnson, "Unit disk graphs," Discrete Mathematics, vol.86, issues 1-3, pp.165–177, 1990.

[9] T. Ohta, M. Fujimoto, R. Oda, and Y. Kakuda, "A class of hierarchical routing protocols based on autonomous clustering for large mobile ad hoc networks," IEICE Trans. Commun., vol.E87-B, no.9, pp.2500-2510, Sept. 2004.