Distributed and Weighted Clustering based on d-Hop Dominating Set for Vehicular Networks

Yan Shi[1*], Xiang Xu[1], Changkai Lu[1] and Shanzhi Chen[2,3]
[1] State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications
Beijing 100876 - China
[e-mail: shiyan@bupt.edu.cn, haoxuxiang@qq.com, frienddeer@163.com]
[2] State Key Laboratory of Wireless Mobile Communications, China Academy of Telecommunications Technology
Beijing 100083 - China
[3] Datang Telecom Technology & Industry Group
Beijing 100083 - China
[e-mail: chensz@datanggroup.cn]
*Corresponding author: Yan Shi

Abstract

Clustering is one of the key technologies in vehicular networks. Constructing and maintaining stable clusters is a challenging task in high mobility environments. DWCM (Distributed and Weighted Clustering based on Mobility Metrics) is proposed in this paper based on the d-hop dominating set of the network. Each vehicle is assigned a priority that describes the cluster relationship. The cluster structure is determined according to the d-hop dominating set, where the vehicles in the d-hop dominating set act as the cluster head nodes. In addition, cluster maintenance handles the cluster structure changes caused by node mobility. The rationality of the proposed algorithm is proven. Simulation results in the NS-2 and VanetMobiSim integrated environment demonstrate the performance advantages.

Keywords: Clustering, d-hop dominating set, priority, mobility metrics, vehicular network

1. Introduction

Vehicular networks are one of the most promising extensions in future wireless and mobile communication systems, aiming to provide efficient V2V (Vehicle-to-Vehicle) and V2I (Vehicle-to-Infrastructure) communications for road safety, traffic efficiency, and infotainment applications [1][2]. Vehicles can use direct, multi-hop, or cluster-based transmission strategies to exchange data with other vehicles or the network infrastructure access points [3].

In a cluster-based transmission strategy, vehicles are organized into multiple clusters. A representative (i.e., a cluster head) is selected for each cluster. The cluster head receives data packets from its cluster members and then relays the packets outside (and vice versa) [3]. For V2I communication, the clustering strategy can efficiently decrease request/data congestion at the access points in the network infrastructure [4]. For V2V communication, the clustering strategy can provide a layered architecture to enhance the network scalability [5][6].

Clustering for MANETs (Mobile Ad-hoc Networks) has attracted considerable research interest [7]. However, the clustering schemes in traditional MANETs may not be suitable for VANETs (Vehicular Ad-hoc Networks) because of their specific requirements [6]. Some of VANETs' notable features should be considered, such as high mobility, the variable density of the VANET nodes, and temporal and spatial dependency in mobility metrics, etc. [1][6]. Constructing and maintaining stable clusters in high mobility environments is crucial for communication continuity. Therefore, clustering is indeed an important aspect in a vehicular network.

The motivation of this paper originates from the following two aspects. First, most existing schemes focus on 1-hop clustering; however, d-hop clustering is capable of providing improved and reliable performance for VANETs with better cluster stability and lower cluster dynamics [6]. Second, mobility-based cluster head selection methods are considered crucial for clustering in VANETs [6] because of their resilience in handling the nodes' mobility variation [7].

Based on the above observations, a d-hop dominating set based clustering algorithm, DWCM (Distributed and Weighted Clustering based on Mobility Metrics), is proposed in this paper. The goal of DWCM is to provide an optimal clustering scheme, adaptive to group mobility feature and with enhanced stability. To catch the group mobility feature, each vehicle is weighted with a priority that defines the cluster relationship between this vehicle and its neighbors.

Based on the d-hop dominating set problem in graph theory, a distributed approach for cluster formation and cluster head selection is designed where vehicles in the d-hop dominating set are selected as the cluster head nodes. In addition, cluster maintenance handles the cluster structure changes caused by node mobility; thus, maintaining cluster stability without incurring tremendous overhead. Simulations are conducted in the NS-2 and VanetMobiSim integrated environment. DWCM presents high stability in cluster head lifetime and re-affiliation times, as well as high scalability in the number of clusters.

The rest of this paper is organized as follows. Related work is reviewed in Section 2. Section 3 describes the network model with assumptions. Section 4 introduces the proposed DWCM clustering approach in detail. Section 5 illustrates the performance using simulation results, and is followed by conclusions in Section 6.

2. Related Work

Clustering schemes in traditional MANETs might not be suitable for VANETs because of VANETs' unique characteristics [6]. The following features should be considered when designing clustering schemes [1][6]:

(1) Vehicular networks show obvious high mobility feature, which incurs frequent topology changes and makes stability the primary design objective in clustering.

(2) Traffic in VANETs presents unique temporal and spatial distributions, rather than the random distributions often used in MANETs. Different traffic states result in a variable vehicle density, e.g., a sparse distribution of vehicles in suburban areas and a dense distribution in congested areas.

(3) Adjacent vehicles show spatial dependency in their mobility metrics, i.e., consistency in the moving direction and similarity in velocity and acceleration. This is because a vehicle's movement pattern can be influenced by, and thus correlated with, vehicles in its neighborhood.

In recent years, more effort has been applied toward efficient clustering in vehicular networks. Based on the distance in hops from an ordinary cluster member node to its cluster head, clustering approaches can be classified as 1-hop clustering or d-hop clustering. Most existing approaches are designed to form 1-hop clusters [9–11]. Only a few considered d-hop clustering algorithms [12–16]. In fact, multi-hop communication is undoubtedly a prevalent scenario in vehicular networks, which makes d-hop clustering a natural choice, with better cluster stability and low cluster dynamics [6].

Cluster size is another factor that should be considered. In [9], the authors defined the cluster size measured in geographical distance between a cluster member node and the cluster head node for better radio efficiency and throughput. In d-hop clustering, the cluster size can be defined in hops from a cluster member node to its cluster head. Such a cluster size should be determined by considering two factors. On one hand, a larger cluster size means fewer clusters, which can reduce network maintenance and control overhead. On the other hand, a network diameter limitation in hops [17] should be considered for efficient multi-hop communication. In addition, clustering that is adaptive to node mobility patterns is believed to have better stability.

Several d-hop clustering schemes in MANETs have also been proposed [12][18–21]. However, the schemes in [18][19] did not consider node mobility patterns. In [20], a load-balancing clustering scheme that focuses on cluster maintenance was proposed, with the goal of limiting the number of mobile nodes in each cluster so the clusters have similar sizes. Although [12][21] had mobility-aware features, the special mobility characteristics in VANETs were not considered. In VANETs—aside from simple mobility metrics, such as moving direction, velocity, and acceleration—more abstract metrics, e.g., mobility dependency, should be defined for an efficient clustering approach [22]. However, existing works lack conformance with the inherent nature of vehicular networks, such as multi-hop communications and mobility dependency between vehicles. Designing a d-hop clustering algorithm that explores the natural group mobility patterns presented by spatial dependency is a meaningful objective.

Existing schemes use different methods for selecting cluster heads [12][23–28]. Methods based on node characteristics use particular node features for cluster head selection, e.g., node ID [23], node degree [24], or use the bus node as the cluster head. In mobility-based methods, some mobility and position information is used for cluster head selection. In prediction-based methods, cluster head selection is often performed based on a prediction of link duration.

Among these different schemes, mobility-based methods are considered key for clustering in VANETs [6]. For example, relative mobility deduced from the received signal strength [12][25], vehicle velocity and position acquired from GPS (Global Positioning System) devices [26][27], and destination information acquired from GPS devices [28] can be used in cluster head selection. Mobility-based methods are considered to be most appropriate for VANETs because of their resilience in handling node mobility variation [8].

3. Network Model and Assumptions

In this paper, the topology of a vehicular network is modeled as an undirected graph $G = (V, E)$, where V is the set of vertexes, and each vertex represents a vehicle in the network. If two vehicles are in direct transmission range of each other, there is an edge between them. Thus, $E \subseteq V \times V$ is the set of edges, and each edge represents a link between vehicles. DWCM depends on the d-hop dominating set to form d-hop clusters. Each vertex is either in the dominating set, or has a path of at most d hops to a vertex in the dominating set. The vertexes in the d-hop dominating set act as cluster heads, whereas the others behave as cluster member nodes.

In this study, the application of the graph theory DS (Dominating Set) problem is motivated by [29–31], where the DS problem provides a promising solution for topology control in MANETs for constructing an efficient network backbone. However, these approaches cannot be applied directly to clustering in vehicular networks because of their different requirements and assumptions.

Topology control in MANETs has unique requirements. For example, the dominating set is often required to be connected and have minimum nodes in the backbone or the shortest path between dominators; thus, the minimum or approximate minimal CDS (Connected Dominating Set) problem is a good choice for describing such approaches. If d-hop clusters are required, the problem evolves to d-hop CDS. In addition, when expecting load balancing or path redundancy features, a k-DS (k-Dominating Set) should be found, where every vertex not in DS has at least k neighboring vertexes in DS.

With regard to optimization objectives, some important factors in MANETs are unsuitable for vehicular networks. For example, energy efficiency is a critical issue for MANETs but meaningless for vehicular networks because the vehicles' batteries recharge during their journey. In addition, vehicle mobility characteristics, along with location information conveniently acquired from GPS devices, should be considered for cluster head selection.

For convenience, the notations used in the rest of this paper are summarized in **Table 1**.

Table 1. Notations

Notation	Description
i.id	Identification of node i
i.pri	Priority of node i
i.state	State of node i, e.g., cluster head, cluster member, gateway, non-clustered
i.up	Upstream node of node i, i.e., the next node along the path from i to its cluster head
i.CH	ID of node i's cluster head
i.hopToCH	Hops from node i to its cluster head node

N_i^d	Set of d-hop neighbors of node i ; the <*id, state, pri*> of each node is recorded
DS_d	d-hop dominating set

4. Proposed DWCM Clustering Algorithm

4.1 d-Hop Dominating Set

As mentioned in Section 3, the network topology is modeled as a graph $G = (V, E)$, and the d-hop dominating set problem in graph theory is adopted here as a basic approach. Formally, the problem can be defined as:

Definition 1 (d-hop neighborhood): The d-hop neighborhood of vertex i is the set of all vertexes within $d (d > 1)$ hops from vertex i .

Definition 2 (d-hop dominating set): For a graph G , a set of vertexes S is called a d-hop dominating set if every vertex is either in S or in the d-hop neighborhood of a vertex in S .

Previous studies on MANET topology control have proven that finding the minimum d-hop dominating set in such a network topology is an NP-complete problem [29–30, 32]. Therefore, suboptimal solution which is the calculation of the minimal d-hop dominating set, should be used to approximate the minimum dominating set problem. In addition, because every node can only acquire the local topology information in vehicular networks, a distributed solution is a natural choice instead of a centralized solution. Based on the above observations, a distributed solution for a d-hop dominating set, based on local topology information, is proposed in this paper. The detailed algorithm is introduced in Section 4.2.

In addition to finding the d-hop dominating set, another important problem is defining the principle in dominating node selection. Each node is assigned a priority, used to evaluate the suitability of a node to act as a cluster head. Depending on different optimization objectives and network environments, the factors involved in the priority calculation could be node identification [19], node degree, residual battery power [32], travel time, node speed or speed deviation [9], and even the integrated result of multiple factors based on a weighted sum or other methods.

As mentioned before, stability is the primary objective in clustering research. Such stability is the result of either less mobility or group mobility (a node and all its neighboring nodes moving in the same direction at approximately the same speed) [32]. For vehicular networks with high mobility, it is undoubtedly a good choice to explore the priority definition that can well describe the group mobility relationship of a node with its surrounding neighbors. Therefore, the node priority is defined as a cluster relationship in this paper. A detailed introduction is given in Section 4.2.

4.2 Cluster Formation and Cluster Head Selection

Cluster formation is the procedure for finding the d-hop dominating set, where the nodes in the dominating set will be selected as the cluster heads. The rules for cluster head selection are introduced first with a correctness proof. The node priority definition is presented thereafter, followed by a distributed approach for cluster formation and cluster head selection.

4.2.1 Rules for cluster head selection

In the proposed DWCM clustering approach, every node communicates with its neighbors to obtain some necessary information. Using this information, a priority value is calculated for

each node. Thereafter, a node decides to become a cluster head if either of the following criteria is satisfied:

Rule 1: The node has the highest priority in its d-hop neighborhood;

Rule 2: The node has the highest priority in the d-hop neighborhood of another node in its d-hop neighborhood.

Each node uses the above rules to determine its role in clustering. If the node satisfies either rule, it behaves as the cluster head node. Otherwise, it acts as an ordinary cluster member node. After all nodes perform such determination, the d-hop dominating set is found based on $G = (V, E)$ derived from the network topology. The correctness of this approach can be proven as follows.

Theorem 1. The set of cluster heads selected by the DWCM algorithm constructs a d-hop dominating set.

Proof: Given the definition of a d-hop dominating set, a node is either a dominator itself (i.e., in the dominating set) or is a d-hop neighbor of a dominating node. This means, for every node i, it satisfies either:

$$i \in DS_d \qquad\qquad (Condition\ 1)$$

or

$$\exists j \,|\, \{i \in N_j^d \quad and \quad j \in DS_d\} \qquad\qquad (Condition\ 2)$$

According to the two rules used by a node to determine whether it is in the d-hop dominating set, the following two cases should be considered:

Case 1: When a node i has the highest priority in its d-hop neighborhood (i.e., i satisfies *Rule 1*), it becomes a cluster head because of *Rule 1*, and acts as a dominator, i.e., $i \in DS_d$. At this time, node i belongs to the d-hop dominating set and satisfies *Condition 1*. The nodes in the d-hop neighborhood of node i, ND_i, satisfy *Condition 2*.

Case 2: For other nodes, node j, $j \in \overline{\bigcup_{i \in DS_d} N_i^d \bigcup DS_d}$, can select the node with the highest priority in its d-hop neighborhood as the cluster head based on *Rule 2*. At this time, the selected cluster head node belongs to the d-hop dominating set and satisfies *Condition 1*. Node j is in the d-hop neighborhood of the selected cluster head, and thus satisfies *Condition 2*. Obviously, a conclusion can be reached that a d-hop dominating set is found after the cluster formation procedure is performed. □

4.2.2 Node priority definition

As mentioned in Section 4.1, when defining node priority, we hope to use the spatial dependency in the mobility metrics between adjacent vehicles. Such a priority describes the group mobility characteristics and natural cluster relationships, and thus becomes a reasonable choice for the criteria in cluster head selection.

Based on our previous work in [33], node priority is defined as the cluster relationship between this node and its neighbors, which is derived from some basic mobility metrics. Given that vehicles can obtain location information conveniently through GPS devices, such a location is represented by its Cartesian coordinates at every time interval T_0. Then, the linear displacement of node i over T_0 can be denoted as:

$$D_{i,T_0} = \sqrt{\Delta x_{T_0}^2 + \Delta y_{T_0}^2} \qquad\qquad (1)$$

where Δx_{T_0} and Δy_{T_0} denote the increase in linear distance of the X and Y coordinates.

Thus, the average velocity of node i over T_0 can be calculated as:

$$v_{i,T_0} = D_{i,T_0} / T_0 \tag{2}$$

Similarly, the average acceleration of node i over T_0 can be defined as:

$$a_{i,T_0} = \Delta v / T_0 \tag{3}$$

where Δv is the velocity variation over T_0.

Based on the above basic mobility metrics, the relative mobility metrics between two adjacent vehicles can be defined. For example, the relative velocity of nodes i and j is defined as:

$$RV(i, j) = 1 - \frac{|v_i - v_j|}{v_{max}} \tag{4}$$

where v_{max} is the vehicles' maximum speed or the upper speed limit of the road.

Similarly, the relative acceleration of nodes i and j is defined as:

$$RA(i, j) = 1 - \frac{|a_i - a_j|}{a_{max}} \tag{5}$$

where a_{max} is the vehicles' maximum acceleration.

Then, the SD (Spatial Dependency) of nodes i and j can be defined considering both relative velocity and acceleration, as shown in Equation (6). The inclusion of the relative acceleration is necessary for a better description of the future relative position of the two nodes.

$$SD(i, j) = RV(i, j) \times RA(i, j) \tag{6}$$

For a vehicle i with n neighbors, its CR (Cluster Relationship) can be defined as its average total spatial dependency on all its n neighbors, and calculated as:

$$CR_i = \frac{\sum_{j=1}^{n} SD(i, j)}{n} \tag{7}$$

Based on the above analysis, a node priority is defined using its *CR*, and then used in the clustering formation and cluster head selection algorithm. A detailed communication procedure for performing the distributed approach is introduced in Section 4.2.3.

4.2.3 Distributed approach

The distributed approach for cluster formation and cluster head selection is illustrated in Algorithm 1. For each node u, Algorithm 1 is executed to determine its state in the cluster, with *ClusterMaxHop* as the parameter to indicate the maximum cluster radius in hops. Each node is initialized to the *Non_Clustered* state. Each node maintains two lists for its d-hop neighbors: (1) *mobilityInfoList*, which contains the mobility information of its d-hop neighbors; (2) *maxPriIdList*, which records <*node id*, *maxPriId*> for every node in its d-hop neighborhood, where *maxPriId* represents the node ID of the neighbor with the highest priority in *node id*'s d-hop neighborhood.

Initially, *mobilityInfoList* and *maxPriIdList* are empty. Then, each node exchanges mobility information within its d-hop neighborhood to fill the *mobilityInfoList*, calculates its priority value accordingly, and exchanges priorities within its d-hop neighborhood. In the period of flooding *MaxPriId*, each node sends its *<node id, maxPriId>* within its d-hop neighborhood and fills in the *maxPriIdList* accordingly. Then, the node can use *maxPriIdList* to determine whether it satisfies *Rule 1* or *Rule 2*. If so, the node can be nominated as the cluster head.

Algorithm 1 Cluster Formation (*ClusterMaxHop*)

```
 1: mobilityInfoList ← ∅ ;
 2: maxPriIdList ← ∅ ;
 3: if period(u) = = Collect_Information then
 4:     SendMobilityInfo(ClusterMaxHop);
 5:     Fill(mobilityInfoList);
 6: else if period(u) = = Calculate_Priority then
 7:     CalculatePriority(mobilityInfoList);
 8:     SendPriDhop(ClusterMaxHop);
 9: else if period(u) = = Flooding_MaxPriId then
10:     SendMaxPriId(ClusterMaxHop);
11:     Fill(maxPriIdList);
12:     if u.id in maxPriIdlist.maxPriId then
13:         u.state = ClusterHead;
14:         for all id in maxPriIdList.id do
15:             if maxPriIdList[id].maxPriId = = u.id then
16:                 SendInvite();
17:             end if
18:         end for
19:     end if
20: end if
```

4.3 Clustering Maintenance

After cluster formation, the cluster structure may suffer frequent changes because of node mobility, e.g., nodes joining or leaving the cluster or a cluster head leaving or failing. Cluster maintenance is another important clustering procedure for handling cluster structure changes, aiming to maintain cluster stability without incurring tremendous overhead.

Some studies have addressed clustering maintenance simply by re-clustering. Obviously, in re-clustering, the cluster head re-election and the necessary information exchange between vehicles result in high computation costs and communication overhead. Therefore, cluster maintenance in DWCM follows a principle similar to [12] and [34], where each node continuously senses the surrounding topology and reacts accordingly with necessary adjustments based on the existing cluster structure.

The detailed cluster maintenance operations are shown in Algorithm 2. Each node executes Algorithm 2 to adjust its state in the cluster to adapt to topology changes. DWCM can address the following maintenance situations:

- Cluster head contention: Node u, nominated as a cluster head through Algorithm 1, performs this operation to contend with another cluster head v, when u and v are within d-hop distance. The node with higher priority wins the contention and acts as

the cluster head, whereas the other gives up its role as cluster head and changes its state to *Non_Clustered*.

- Cluster gateway discovery: If a cluster member can hear messages from more than one cluster head, it behaves as a cluster gateway and changes its state accordingly.

Algorithm 2 Cluster Maintenance

1: N_u^d records *u*'s d-hop neighborhood information (<*id, status, priority*> for each neighbor);
2: **if** *u.status* $==$ *ClusterHead* **then**
3:      **if** (node *v* in N_u^d) and (*v.state* $==$ *ClusterHead*) **then**
4:          *wait* (*Cluster_Contention_Interval*);
5:          **if** (*v.id* still in N_u^d) and (*u.pri* $<$ *v.pri*) **then**
6:              *u.state* = *Non_Clustered*;
7:          **Endif**
8:      **Endif**
9:      **if** lost all clusterMembers **then**
10:          *u.state* = *Non_Clustered*;
11:      **Endif**
12: **Endif**
13: **if** *u.state* $==$ *ClusterMember* **then**
14:      **if** (node *v* in N_u^d) and (*v.state* $==$ *ClusterHead*) and (*v.id* $\neq$ *u.CH*) **then**
15:          *u.state* = *ClusterGateway*;
16:      **Endif**
17: **Endif**
18: **if** (*u.state* $==$ *ClusterMember*) or (*u.state* $==$ *ClusterGateway*) **then**
19:      **if** cannot hear from the ClusterHead and cannot hear from its upStreamNode during Cluster_Head_Timeout_Interval **then**
20:          *u.state* = *Non_Clustered*;
21:      **Endif**
22: **Endif**
23: **if** (*u.state* $==$ *Non_Clustered*) **then**
24:      **if** neighbors are all nonclustered nodes **then**
25:          *Cluster_Formation*();
26:      **Else**
27:          *N1* $\subseteq N_u^d$ is the set of *u*'s neighbors who have joined a Cluster;
28:          **for all** node in *N1* **do**
29:              node *v*←the node with *max(SD)*;
30:              **if** (*SD(u, v)* $>$ a predefined threshold *Cluster_Min_SD*) and the direction of *u, v* are almost the same and (*Hop(v)* $<$ *ClusterMaxHop*) **then**
31:                  *u.state* = *ClusterMember*;
32:                  *Hop(u)* = *Hop(v)*+1;
33:                  **if** *v.state* $==$ *ClusterMember* **then**
34:                      *upStreamNode(u)* = *v*;
35:                  **Endif**
36:              **Else**
37:                  *N1* = *N1*−*v*;
38:              **Endif**
39:          **Endfor**
40:      **Endif**
41: **Endif**

- Isolated node discovery: If a cluster head loses all its cluster members, or a cluster member or gateway does not hear from its cluster head and its upstream node during a predefined interval, it becomes an isolated node. Then, its state should be changed to *Non_Clustered*.

- Joining a new cluster: A *Non_Clustered* node u selects a suitable clustered neighbor v and joins v's cluster. The criteria for selecting v include: 1) node v has the highest SD value in u's neighborhood, and this value is higher than a predefined threshold; 2) the moving direction of u and v is almost the same, i.e., below a threshold. This is used to avoid joining a cluster that moves in the opposite direction; and 3) the distance from u to v is less than the *ClusterMaxHop* constraint.

5. Simulation Results

Simulations were conducted in the network simulator NS-2 [35] to verify the effectiveness and performance of the proposed DWCM clustering algorithm. Because there are no predefined mobility models for vehicle movement, VanetMobiSim [36] was used as the microscopic traffic generator to provide traffic simulation data that describe realistic vehicle traffic. A common framework was designed in NS-2 to implement the clustering algorithms. The simulation parameters are defined in **Table 2**.

Table 2. Simulation Parameters

Parameter	Description	Value
N	Number of nodes in simulation	[50, 175]
$m \times n$	Network size	$10000\ m \times 6000\ m$
S	Max speed of node movement	[10 m/s, 30 m/s]
T_r	Transmission range of vehicles	[50 m, 300 m]
BI	Broadcast interval	$1\ s \pm \Delta, \Delta \in [0, \frac{1}{5}\ s]$
T	Simulation time	2500 s
M	Mobility model	IDM-LC

The proposed DWCM clustering algorithm and some classical approaches (including Lowest-ID, Highest-degree, MOBIC, and MobDHop), were implemented in NS-2. Lowest-ID [23] selects cluster heads with the lowest node ID and easily causes re-clustering when an un-clustered node with a lower ID reaches a cluster head's range. Highest-degree [24] uses the local highest node degree as the attribute for cluster head selection. However, it still has the same problem as Lowest-ID, where re-clustering occurs frequently.

MOBIC [25] proposes an aggregate local mobility metric for the cluster formation process, such that mobile nodes with a lower speed than their neighbors have the chance to become cluster heads. MobDHop [12] forms variable diameter clusters based on the node mobility pattern in MANETs. It introduces a new metric for measuring the distance variation between nodes over time to estimate the relative mobility of two nodes. The multi-hop clustering schemes of MobDHop and the proposed DWCM are evaluated with 2-hop performance (i.e., d = 2) as a tradeoff between low cluster dynamics and complexity. It is practical to acquire 2-hop information with acceptable complexity, while providing enhanced stability compared with 1-hop clustering.

These clustering approaches were simulated under a series of similar configurations. For performance evaluation, the following performance metrics were chosen:

- Mean cluster head lifetime: Also known as cluster lifetime, this refers to the average time duration of all the nodes acting as cluster heads during the entire simulation. This is an important metric for cluster stability.

- Mean re-affiliation times: The mean times of a node (cluster head or member) leaving one cluster and joining another. A lower value reduces bandwidth consumption during message exchanges for member updates.

- Number of clusters: The total number of clusters formed in the network during simulation time. This is a metric for the scalability of the clustering approaches.

Firstly, the performance of Lowest-ID based 1-hop clustering, Highest-degree based 1-hop clustering, MOBIC based 1-hop clustering, MobDHop based 2-hop clustering, and DWCM based 2-hop clustering was simulated when the transmission range of the vehicles varied from 50 m to 300 m with a fixed N = 125 and S = 20 m/s. **Fig. 1** shows the cluster stability of the above approaches when varying the transmission range. From the results, we can see that DWCM outperforms the other approaches in both cluster head lifetime and re-affiliation times.
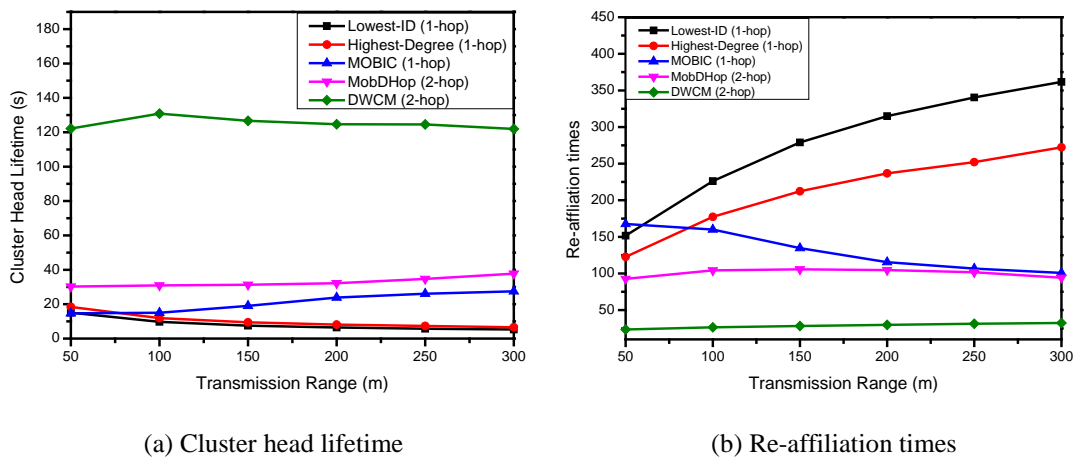

(a) Cluster head lifetime   (b) Re-affiliation times
Fig. 1. Cluster head lifetime and re-affiliation times vs. transmission range

As shown in **Fig. 1(a)**, in terms of mean cluster head lifetime, DWCM, MobDHop, and MOBIC show obvious performance advantages compared with Lowest-ID and Highest-Degree. This is because DWCM, MobDHop, and MOBIC use relative mobility metric related parameters in cluster formation and cluster head selection. Such parameters can describe the natural group mobility feature and help select more suitable cluster head nodes; thus, improving the cluster lifetime.

Both MobDHop and MOBIC estimate the distance from a node to its neighbor based on the measured received signal strength from that particular neighbor. MobDHop performs better than MOBIC because it uses the standard deviation of the distance variation based on a series of continuous measurements to capture the group mobility, whereas MOBIC only uses a pair of consecutive measurements. DWCM can achieve better performance because more relative mobility metrics, e.g., relative direction, relative velocity, and relative acceleration, are used to define the cluster relationship. **Fig. 1(b)** shows that DWCM performs obviously better than the other approaches in terms of mean re-affiliation times because it considers the relative moving direction between different nodes in cluster formation and maintenance operations; thus, it avoids joining a cluster that moves in the opposite direction.

Then, the performance of Lowest-ID based 1-hop clustering, Highest-degree based 1-hop clustering, MOBIC based 1-hop clustering, MobDHop based 2-hop clustering, and DWCM based 2-hop clustering was simulated when the maximum speed of the vehicles varied from 10 m/s to 30 m/s with a fixed N = 125 and T_r = 200 m. Such simulations are used to investigate the effect of vehicle speed on cluster performance. **Fig. 2** illustrates the mean cluster head lifetime and mean re-affiliation times with varying maximum vehicle speed in **Figs. 2(a)** and **2(b),** respectively. We can see performance results similar to those shown in **Fig. 1**.

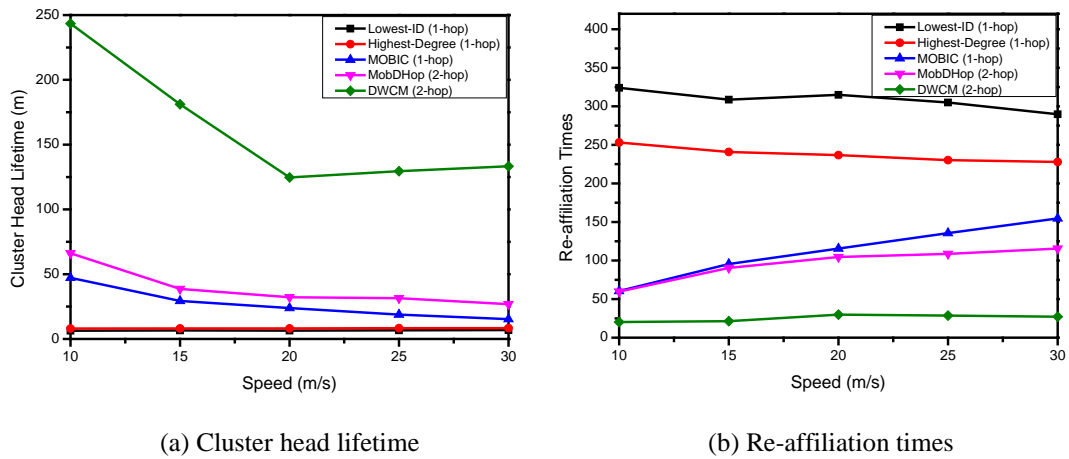(a) Cluster head lifetime   (b) Re-affiliation times
Fig. 2. Cluster head lifetime and re-affiliation times vs. speed

Figs. 3(a) and **3(b)** show the stability performance of Lowest-ID based 1-hop clustering, Highest-degree based 1-hop clustering, MOBIC based 1-hop clustering, MobDHop based 2-hop clustering, and DWCM based 2-hop clustering when the number of vehicles varies from 50 to 75 with a fixed S = 20 m/s and T_r = 200 m. The clustering algorithms that consider mobility metrics reveal more obvious performance advantages as the number of vehicles increases. This is because the spatial dependency between the vehicles rises with the number of vehicles, thus improving the cluster stability.
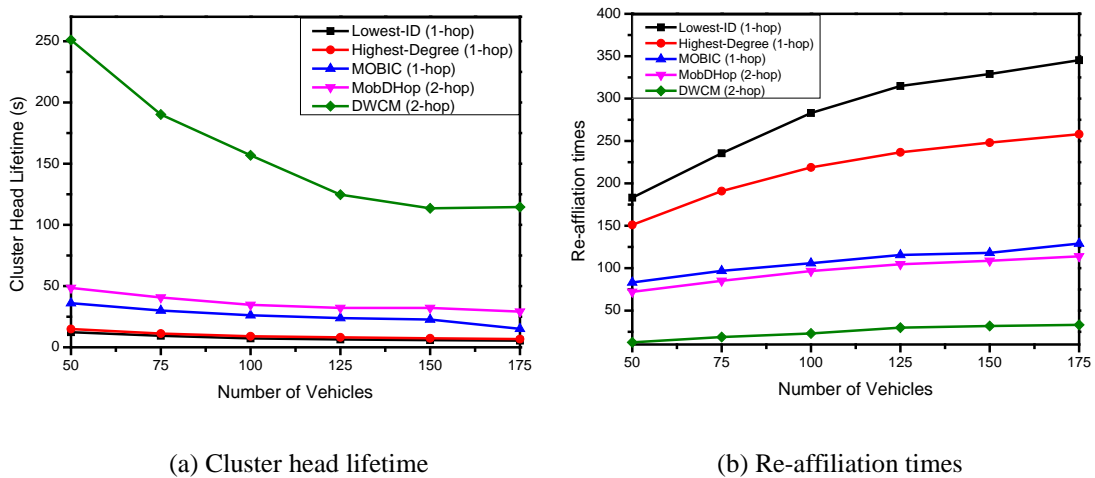

(a) Cluster head lifetime   (b) Re-affiliation times
Fig. 3. Cluster head lifetime and re-affiliation times vs. the number of vehicles

Figs. 4(a), **4(b),** and **4(c)** illustrate the cluster numbers of different approaches for various values of transmission range, vehicle speed, and number of vehicles. These figures show that multi-hop clustering approaches (e.g., MobDHop and DWCM) form obviously fewer clusters than 1-hop clustering approaches (e.g., Lowest-ID, Highest-Degree, and MOBIC) in the simulation. A smaller cluster number is desirable because the delay and overhead can be reduced in cluster-based hierarchical routing. Thus, multi-hop clustering is a more reasonable choice for large-scale vehicular networks.

(a) Varying transmission range   (b) Varying max speed

(c) Varying number of vehicles
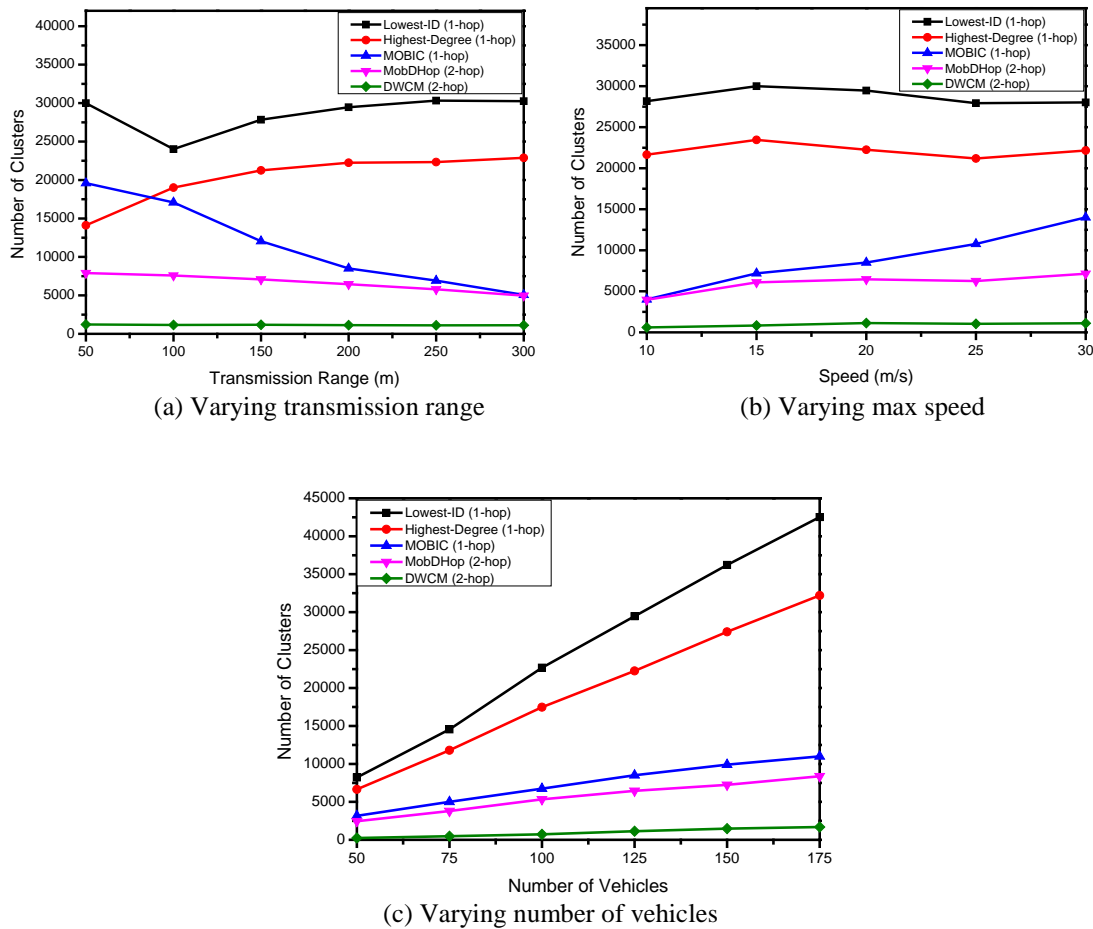
Fig. 4. Cluster number when the transmission range, speed, and number of vehicles vary

We also note the difference between the simulation results in this paper and the performance data in other research papers. It seems that the same algorithm shows worse performance in our simulation. Such a deviation is caused by the mobility models used in the simulation. In this study, VanetMobiSim was used as the microscopic traffic generator to provide more realistic traffic simulation data. The IDM-LC (Intelligent Driver Model with Lane Change) model was used in the simulations. IDM-LC has the functions of IM (Intersection Management) and LC (Lane Change). The IM function induces congestion at intersections. Therefore, the cluster heads and members are more likely to disconnect from each other; thus, resulting in more cluster dismissals, joining new clusters, and re-clustering. The cluster head lifetime decreases, whereas the re-affiliation times increase accordingly. The LC function permits overtaking behaviors, which decrease the mobility dependency between

vehicles, and thus degrade the performance of clustering algorithms based on mobility metrics.

Experiments were also conducted to investigate the effects of the cluster radius in hops (i.e., the value of d) on clustering performance. **Figs. 5(a), 5(b),** and **5(c)** show the clustering performance of DWCM for mean cluster head lifetime, mean re-affiliation times, and number of clusters, respectively, when $d = 2$, 3, and 4 hops with a fixed $N = 125$ and $T_r = 200$ m, with the maximum speed varying from 10 m/s to 30 m/s. Although a clustering performance improvement (i.e., increase in cluster head lifetime, decrease in re-affiliation times, and decrease in number of clusters) can be observed when we increase the cluster radius, the performance benefit is rather limited and will shrink with a larger d value.
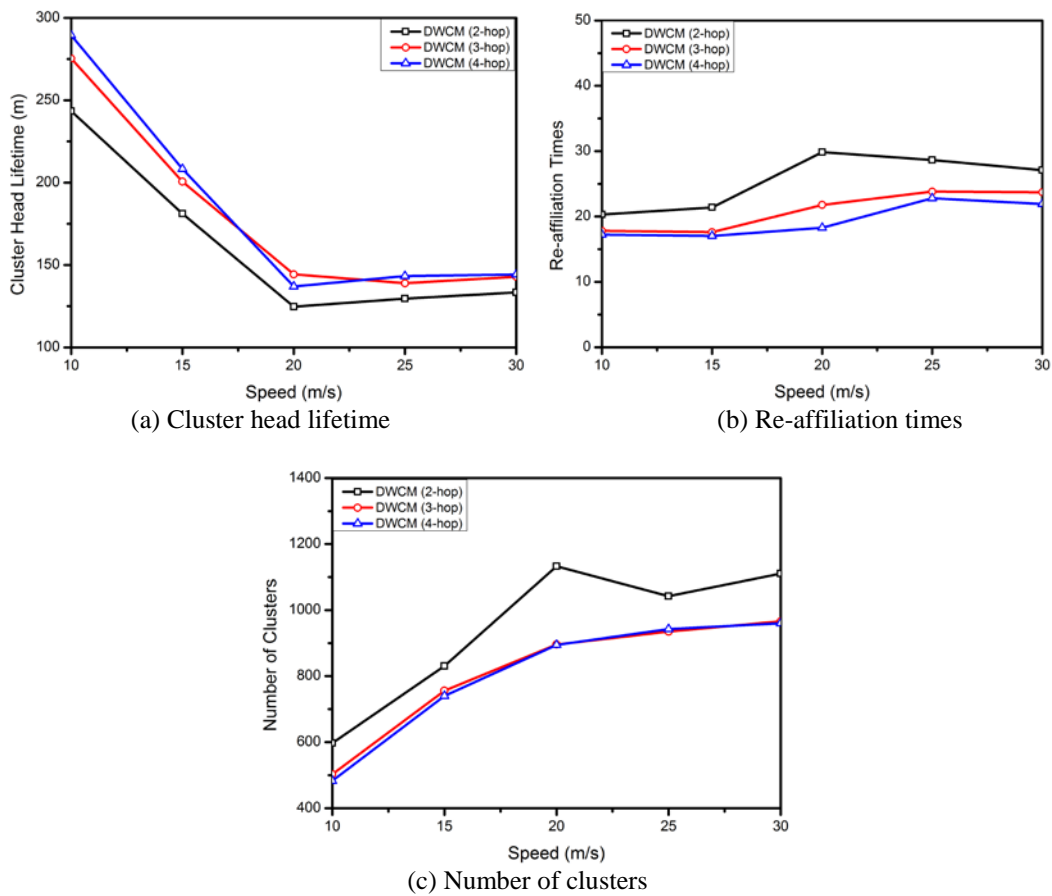

(a) Cluster head lifetime   (b) Re-affiliation times

(c) Number of clusters

Fig. 5. Effects of d value on clustering performance

As the optimality of a clustering algorithm is defined as being able to form as few stable clusters as possible at a reasonable overhead [12], the overhead introduced by a larger d value should also be considered. For DWCM, the message overhead will increase with a larger number of hops because the vehicles' mobility information is disseminated within the d-hop range. The message overhead of the cluster formation procedure in the worst case can be as high as $O(N^d)$. When entering the cluster maintenance procedure, the message overhead for each topology change can be described as $O(m \cdot d)$, where m is the average number of members in a cluster. Therefore, for practical feasibility, the choice of d should consider the tradeoff between low cluster dynamics and overhead.

In addition, when considering practical applications, different city and rural scenarios can affect clustering performance. As observed in the simulation, cluster stability is affected by vehicle density, speed, and wireless transmission range. In city scenarios, the vehicle density is higher, whereas the speed is lower, compared with rural scenarios. In addition, the node priority definition in this paper is based on the group mobility features in VANETs. The spatial dependency in mobility metrics between adjacent vehicles in one road segment is more obvious in city scenarios than in rural ones. Therefore, in general, DWCM will perform better in city scenarios than in rural scenarios.

However, DWCM performance in city scenarios might vary depending on the different road layouts. For example, in a scenario with a dense deployment of intersections and traffic signals, the dynamics of the clusters increase because vehicles often join or leave the clusters; thus, incurring large cluster maintenance overhead. Another case is city expressways, where vehicles move within a long segment to maintain the stable group mobility feature. DWCM is expected to have better performance in such a scenario.

6. Conclusions and Future Work

In this study, DWCM, a distributed and weighted d-hop clustering method based on mobility metrics, was proposed. The goal of DWCM is to construct and maintain stable multi-hop clusters in vehicular networks. A weighted undirected graph was used as the network model. Each vertex in this graph was assigned a priority that described the group mobility feature in the vehicular network. A d-hop dominating set was found for cluster head nomination and the correctness of this algorithm was proven. In addition, cluster maintenance was used to handle the cluster structure changes, including cluster head contention, cluster gateway discovery, isolated node discovery, and joining a new cluster. The simulation results in the NS-2 and VanetMobiSim integrated environment showed that DWCM outperformed other classical clustering approaches in terms of cluster stability, with longer cluster head lifetimes and fewer re-affiliation times. In addition, d-hop clustering in DWCM forms a smaller number of clusters with high scalability. It is notable that recent work [37] has explored the topology characteristics based on large-scale realistic vehicle mobility traces. A good clustering algorithm should be capable of identifying and adapting to the inherent group mobility patterns of a temporally evolving topology. Designing such a clustering algorithm by employing the complex network theory and statistical physics theory will be our ongoing and future work.

References

[1] S. Al-Sultan, M. M. Al-Doori, A. H. Al-Bayatti and H. A. Zedan, "A comprehensive survey on vehicular Ad Hoc network," *Journal of Network and Computer Applications*, vol. 37, no. 1, pp. 380-392, January, 2014. Article (CrossRef Link)

[2] G. Karagiannis, O. Altintas, E. Ekici, G. Heijenk, B. Jarupan, K. Lin and T. Weil, "Vehicular networking: a survey and tutorial on requirements, architectures, challenges, standards and solutions," *IEEE Communications Surveys & Tutorials*, vol. 13, no. 4, pp. 584-616, December, 2011. Article (CrossRef Link)

[3] E. Hossain, G. Chow, V. C. M. Leung, R. D. Mcleod, J. Mišić, V. W. S. Wong and O. Yang, "Vehicular telematics over heterogeneous wireless networks: A survey," *Computer Communications*, vol. 33, no. 7, pp. 775-793, May, 2010. Article (CrossRef Link)

[4]   H. Su and X. Zhang, "Clustering-based multichannel MAC protocols for QoS provisionings over vehicular ad hoc networks," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 6, pp. 3309-3323, December, 2007. Article (CrossRef Link)

[5]   C. Konstantopoulos, D. Gavalas and G. Pantziou, "Clustering in mobile ad hoc networks through neighborhood stability-based mobility prediction," *Computer Networks*, vol. 52, no. 9, pp. 1797-1824, June, 2008. Article (CrossRef Link)

[6]   R. S. Bali, N. Kumar and J. J. P. C. Rodrigues, "Clustering in vehicular Ad Hoc networks: taxonomy, challenges and solutions," *Vehicular Communications*, vol. 1, no. 3. pp. 134-152, July, 2014. Article (CrossRef Link)

[7]   J. Y. Yu and P. H. J. Chong, "A survey of clustering schemes for mobile ad hoc networks," *IEEE Communications Surveys & Tutorials*, vol. 7, no. 1, pp. 32-48, May, 2005. Article (CrossRef Link)

[8]   A. Fonseca and T. Vazão, "Applicability of position-based routing for VANET in high-ways and urban environment," *Journal of Network and Computer Applications*, vol. 36, no. 3, pp.961-973, May, 2013. Article (CrossRef Link)

[9]   Z. Wang, L. Liu, M. C. Zhou and N. Ansari, "A position-based clustering technique for Ad Hoc Intervehicle communication," *IEEE Transactions on Systems, Man and Cybernetics - Part C: Applications and Reviews*, vol. 38, no. 2, pp. 201-208, May, 2008. Article (CrossRef Link)

[10] A. Benslimane, T. Taleb and R. Sivaraj, "Dynamic clustering-based adaptive mobile gateway management in integrated VANET-3G heterogeneous wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 3, pp. 559-570, March, 2011. Article (CrossRef Link)

[11] C. Shea, B. Hassanabadi and S. Valaee, "Mobility-based clustering in VANETs using Affinity Propagation," in *Proc. of IEEE Global Telecommunications Conference 2009*, pp. 1-6, November 30-December 4, 2009. Article (CrossRef Link)

[12] I. I. Er and W. K. G. Seah, "Performance analysis of mobility-based d-hop (MobDHop) clustering algorithm for mobile ad hoc networks," *Computer Networks*, vol. 50, no. 17, pp. 3375-3399, December, 2006. Article (CrossRef Link)

[13] Z. Zhang, A. Boukerche and R. W. Pazzi, "A novel multi-hop clustering scheme for vehicular ad-hoc networks," in *Proc. of 9th ACM International Symposium on Mobility Management and Wireless Access*. pp. 19–26, October 31-November 4, 2011. Article (CrossRef Link)

[14] L. Zhang and H. El-Sayed, "A novel cluster-based protocol for topology discovery in vehicular ad hoc network," in *Proc. of ACM 3rd International Conference on Ambient Systems, Networks and Technologies*, pp.525–534, August 27-29, 2012. Article (CrossRef Link)

[15] W. Li, A. Tizghadam and A. Leon-Garcia, "Robust clustering for connected vehicles using local network criticality," in *Proc. of IEEE Conference on Communications 2012*, pp.7157–7161, June 10-15, 2012. Article (CrossRef Link)

[16] E. Dror, A. Chen and Z. Lotker, "Fast randomized algorithm for 2-hops clustering in vehicular ad-hoc networks," *Ad Hoc Networks*, vol. 11, no. 7, pp. 2002-2015, September, 2013. Article (CrossRef Link)

[17] M. Aoki and H. Fuji, "Inter-vehicle communication: technical issues on vehicle control application," *IEEE Communications Magazine*, vol.34, no. 10, pp. 93-99, October, 1996. Article (CrossRef Link)

[18] G. Chen, F. G. Nocetti, J. S. Gonzalez and I. Stojmenovic, "Connectivity based k-hop clustering in wireless networks," in *Proc. of IEEE 35th Annual Hawaii International Conference on System Sciences*, pp. 2450-2459, January 7-10, 2002, Article (CrossRef Link)

[19] A. D. Amis, R. Prakash, T. H. P. Vuong and D. T. Huynh, "Max-min d-cluster formation in wireless ad hoc networks," in *Proc. of IEEE 19th Conference on Computer Communications*, pp. 32-41, March 26-30, 2000. Article (CrossRef Link)

[20] T. Ohta, S. Inoue and Y. Kakuda, "An adaptive multihop clustering scheme for highly mobile ad hoc networks," in *Proc. of 6th International Symposium on Autonomous Decentralized Systems*, pp. 293-300, April 9-11, 2003. Article (CrossRef Link)

[21] A. B. McDonald and T. F. Znati, "Design and performance of a distributed dynamic clustering algorithm for ad-hoc networks," in *Proc. of 34th Annual Simulation Symposium*, pp. 27–35, April 22-26, 2001. Article (CrossRef Link)

[22] Y. Zhang and J. M. Ng, "A distributed group mobility adaptive clustering algorithm for mobile ad hoc networks," in *Proc. of IEEE International Conference on Communications 2008*, pp. 3161-3165, May 19-23, 2008. Article (CrossRef Link)

[23] A. Ephremides, J. E. Wieselthier and D. J. Baker, "A design concept for reliable mobile radio networks with frequency hopping signaling," in *Proc. of the IEEE*, vol. 75, no. 1, pp. 56-73, January, 1987. Article (CrossRef Link)

[24] M. Gerla and J. T. C. Tsai, "Multiuser, Mobile, Multimedia Radio Network," *Wireless Networks*, vol. 1, no. 3, pp. 255-265, September, 1995. Article (CrossRef Link)

[25] P. Basu, N. Khan and T. D. C. Little, "A mobility based metric for clustering in mobile ad hoc networks," in *Proc. of IEEE 21st International Conference on Distributed Computing Systems Workshop*, pp. 413-418, April 16-19, 2001. Article (CrossRef Link)

[26] E. Souza, I. Nikolaidis and P. Gburzynski, "A new aggregate local mobility (ALM) clustering algorithm for VANETs," in *Proc. of IEEE International Conference on Communications 2010*, pp.1-5, May 23-27, 2010. Article (CrossRef Link)

[27] C. Shea, B. Hassanabadi and S. Valaee, "Mobility-based clustering in VANETs using affinity propagation," in *Proc. of IEEE Global Telecommunications Conference 2009*, pp. 1-6, November 30-December 4, 2009. Article (CrossRef Link)

[28] M. M. C. Morales, C. S. Hong and Y. C. Bang, "An adaptable mobility-aware clustering algorithm in vehicular networks," in *Proc. of IEEE 13th Asia-Pacific Network Operations and Management Symposium*, pp. 1-6, September 21-23, 2011. Article (CrossRef Link)

[29] L. Bao and J. J. Garcia-Luna-Aceves, "Topology management in ad hoc networks," in *Proc. of the 4th ACM international symposium on Mobile ad hoc networking & computing*, pp. 129-140, June 1-3, 2003. Article (CrossRef Link)

[30] H. Y. Yang, C. H. Lin and M. J. Tsai, "Distributed algorithm for efficient construction and maintenance of connected k-hop dominating sets in mobile ad-hoc networks," *IEEE Transactions on Mobile Computing*, vol. 7, no. 4, pp. 444-457, April, 2008. Article (CrossRef Link)

[31] W. K. Wolterink, "A content-based routing protocol for mobile ad-hoc networks using a distributed connected k-hop dominating set as a backbone," *Master thesis*, University of Twente, Enschede, Netherlands, 2008.

[32] V. S. Anitha and M. P. Sebastian, "(k,r)-Dominating set-based, weighted and adaptive clustering algorithms for mobile ad hoc networks," *IET Communications*, vol. 5, no. 13, pp. 1836-1853, September, 2011. Article (CrossRef Link)

[33] W. Fan, Y. Shi, S. Z. Chen and L. H. Zou, "A mobility metrics based dynamic clustering algorithm for VANETs," in *Proc. of IET International Conference on Communication Technology and Application 2011*, pp.752-756, October 14-16, 2011. Article (CrossRef Link)

[34] S. Basagni, "Distributed clustering for ad hoc networks," in *Proc. of 4th International Symposium on Parallel Architectures, Algorithms, and Networks*, pp. 1087-4089, June 23-25, 1999. Article (CrossRef Link)

[35] K. Fall and K. Varadhan, "The ns Manual," http://www.isi.edu/nsnam/ns/, 2002.

[36] J. Harri and M. Fiore, "VanetMobiSim - Vehicular Ad hoc Network mobility extension to the CanuMobiSim framework," http://vanet.eurecom.fr/.

[37] D. Naboulsi and M. Fiore, "On the instantaneous topology of a large-scale urban vehicular network: the cologne case," in *Proc. of the 14th ACM international symposium on Mobile ad hoc networking and computing*, pp. 167-176, July 29-August 1, 2013. Article (CrossRef Link)

Yan Shi received her Ph. D. degree from Beijing University of Posts and Telecommunications (BUPT), China, in 2007. She is currently a research staff of the State Key Laboratory of Networking and Switching Technology. Her current research interests include network architecture evolution, protocol design and performance optimization of future networks and mobile computing, especially mobility management technology.

Xiang Xu received his B. Sc. degree from Beihua University in 2014 and is working towards the M. Sc. degree in State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications (BUPT) in China. His current research interests include vehicular networks and mobile Internet.

Changkai Lu received his B. Sc. degree from Shandong Agricultural University in 2011 and is working towards the M. Sc. degree in State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications (BUPT) in China. His current research interests include vehicular networks and mobile Internet.

Shanzhi Chen received his Ph. D. degree from Beijing University of Posts and Telecommunications (BUPT), China, in 1997. He joined Datang Telecom Technology & Industry Group in 1994, and has been serving as CTO since 2008. He was a member of the steering expert group on information technology of the 863 Program of China from 1999 to 2011. He is the director of State Key Laboratory of Wireless Mobile Communications, and the board member of Semiconductor Manufacturing International Corporation (SMIC). He has made great contributions to TD-SCDMA 3G industrialization and TD-LTE-advanced 4G standardization. He received State Science and Technology Progress Award of China in 2001 and 2012. His current research interests include wireless mobile communications, IoT and emergency communications.