

RESEARCH

Open Access

Distributed multi-hop clustering algorithm for VANETs based on neighborhood follow

Yuzhong Chen^{1,2}, Mingyue Fang^{1,2}, Song Shi^{1,2}, Wenzhong Guo^{1,2} and Xianghan Zheng^{1,2*}

Abstract

Vehicular *ad hoc* networks (VANETs) have become important components of metropolitan area networks, and clustering for VANETS provides many advantages. However, the stability of current clustering algorithms exhibits poor robustness because a VANET is a highly dynamic scenario. In this study, a novel multi-hop clustering scheme for VANETS, which generates cluster heads (CHs) via neighborhood follow relationship between vehicles, is proposed. The scheme is based on a reasonable assumption that a vehicle cannot certainly identify which vehicle in its multi-hop neighbors is the most suitable to be its CH, but it can easily grasp which vehicle in one-hop distance is the most stable and similar with it, and thus, they most likely belong to the same cluster. Consequently, a vehicle can choose its CH by following the most stable vehicle. The relative mobility between two vehicles combining the gains based on the followed number and the historical following information enables a vehicle to select which target to follow. Extensive simulation experiments are conducted to validate the performance of the proposed clustering scheme.

Keywords: VANETs; Clustering; Multi-hop; Neighborhood follow

1 Introduction

As a new form of mobile *ad hoc* networks (MANETs), the vehicular *ad hoc* network (VANET) has emerged with the rapid development of radio technology that allows vehicle-to-vehicle communication [1]. VANET is one of the important components of an intelligent transport system because it holds great potential in traffic accident warning, traffic flow control, as well as in providing information services and extra serviceability.

Clustering in VANET exhibits good scalability, because clustering can provide a simple information management mechanism and improve communication efficiency. Therefore, clustering algorithms for VANETs are attracting increasing attention.

Unlike traditional MANETs, VANETs exhibit new features such as rapid movement of vehicles, frequent changing of network topology, and limited driving directions; moreover, it does not consider energy problems [2]. These particular features are the reason why traditional clustering algorithms for MANETs can hardly be applied in VANETs. Cluster stability is an important

requirement in VANETs. Vehicles move fast which makes clusters broken easily and further affects the routing efficiency. Moreover, unstable clusters are prone to generate more control packets in VANETs and make the networks overload.

Considerable research has explored clustering algorithms for VANETs to satisfy the requirements of their new features [3-10]. Most of these studies are based on one-hop clustering, which only allows communication between a cluster member (CM) and its cluster head (CH) with one-hop distance. The coverage of clusters is small in one-hop clustering, which leads to excess CHs and high-maintenance overhead. Consequently, several multi-hop clustering algorithms have been proposed in the past years [11-13]. These algorithms can extend the coverage of clusters, reduce the number of CHs, and improve cluster stability. However, some issues remain in multi-hop clustering for VANETs. For example, cluster stability must be further improved and maintenance cost must be reduced. Thus, comprehensive solutions must be developed.

This study proposes a distributed multi-hop clustering algorithm for VANETs based on neighborhood follow (DMCNF) to enable fast and stable network setup. The main idea of this study is explained as follows. In large-scale and complex VANETs, a vehicle can hardly acquire

* Correspondence: xianghan.zheng@fzu.edu.cn

¹College of Mathematics and Computer Science, Fuzhou University, No. 2 Xue Yuan Road, 350108 Fuzhou, China

²Fujian Key Laboratory of Network Computing and Intelligent Information Processing, No. 2 Xue Yuan Road, 350108 Fuzhou, China

precise details of multi-hop distanced vehicles and can hardly decide which CH to choose among multi-hop neighbors. However, a vehicle can easily master local information and determine which vehicle within a one-hop distance is the most stable/similar to it, and thus, they most likely belong to the same cluster. Consequently, a vehicle can choose its CH by following the most similar vehicle. This mechanism is identified as the neighborhood follow relationship. In turn, selecting CHs becomes a feedback from the neighborhood follow relationship. Accordingly, this study attempts to quickly obtain the aforementioned relationship chain from large-scale VANETs. CHs are then chosen according to the obtained relationship among vehicles, and other vehicles directly or indirectly follow CHs to form clusters. The main contributions of this work are as follows.

- (1) A novel cluster model based on one-hop neighborhood follow is proposed. In the model, a cluster has a CH, which is directly or indirectly followed by other vehicles. The structure of DMCNF can steadily evolve in highly dynamic VANETs.
- (2) A neighborhood follow strategy is introduced for vehicles to choose and follow stable target vehicles from one-hop neighbors. Through this strategy, vehicles can adaptively update neighborhood follow information. Consequently, clusters are formed and maintained.
- (3) Through the neighborhood follow strategy, clusters are formed and maintained in a distributed manner. Vehicles are only required to regularly communicate with its one-hop neighbors for updating the neighborhood following information and maintaining clusters.
- (4) DMCNF does not depend on location service but still provides grouping of related vehicles and fast response to topology changes.

The rest of this paper is organized as follows. Section 2 discusses the review of related literature, and Section 3 provides the preliminaries of the study. Section 4 describes the DMCNF algorithm, and Section 5 presents the experimental results. Finally, Section 6 concludes the paper and suggests a potential subject for future work.

2 Related work

Clustering is a well-known means of organizing networks in MANETs. Many clustering solutions, including identifier neighbor-based clustering, topology-based clustering, mobility-based clustering, energy-based clustering, and weight-based clustering, have been proposed [14-21]. However, these clustering solutions significantly differ from vehicular clustering. MANETs are primarily limited because of their energy [22] and processing power; hence,

their clustering is optimized for low-resource usage. However, vehicles are not only rich in resources, but they are also highly mobile. Consequently, clustering algorithms for MANETs are not effective in VANETs, and new solutions must be developed.

Mobility-based clustering algorithm (MOBIC) [16] is a popular clustering algorithm mentioned in various studies. This approach is based on the lowest-ID algorithm; yet, it uses a signal power level mobility metric that is derived from successive receptions. MOBIC does not scale well in VANETs because it is a simple algorithm designed for MANETs; nevertheless, it is frequently compared with other VANET clustering algorithms.

Hafeez et al. [9] proposed a novel clustering algorithm for VANETs by considering speed as the main influential factor to form clusters. These researchers also attempted to improve cluster stability via the fuzzy processing of speed. The algorithm introduced by Hafeez et al. chooses the second optimal vehicle as the temporary CH when the original one becomes unavailable. The algorithm is applied to high-mobility scenarios, but CHs frequently change when they move fast. The rapid change in network topology induces the unstable performance of temporary CHs, which results in unstable clusters.

The affinity propagation (AP) algorithm is one of the most stable clustering algorithms that have been recently proposed. Applying AP to VANETs [10] remarkably improves cluster stability. Nevertheless, AP is a distance-based clustering algorithm, which results in the frequent changing of CHs when speed changes dramatically. Moreover, AP requires several iterative loops that increase the delay time of cluster formation.

The majority of the other known clustering solutions for VANETs are studied in [8]. Goonewardene et al. [4] presented a novel algorithm with the rare feature of cluster overlapping. In particular, this algorithm considers speed, location, and direction via the Global Positioning System (GPS) or other similar services to form clusters. The aggregate local mobility (ALM) algorithm [5] is a new beacon-based clustering scheme that aims to extend the lifetime of clusters by using ALM to decide cluster reorganization. Meanwhile, Rawashdeh and Mahmud [7] considered speed and relative direction to present a novel speed-overlapped clustering algorithm for highways. This system also depends on location services.

The majority of the proposed clustering algorithms for VANET depend on GPS, which may not be the ideal option [23-25]. Positioning services are not available everywhere, and even if they are, their accuracy can significantly vary. Inaccuracy, even within a few meters, endangers position service-based clustering algorithms because it can severely affect cluster stability and leads to communication failure, which are both unacceptable. The complexity and reliability of the entire clustering

system are also influenced by its dependence on positioning services. Nonetheless, vehicles should be able to communicate even when positioning services are unavailable or unreliable.

Moreover, most of the aforementioned algorithms are based on one-hop clustering, which only allows communication among one-hop neighbors. Consequently, the coverage rank is small and many clusters are formed, which decrease cluster stability. Cluster formation algorithms should be designed to guarantee cluster stability which is crucial to reduce the maintenance cost of clusters and increase throughput and routing efficiency.

Several scholars have recently analyzed multi-hop clustering algorithms, which are rare but exhibit good achievements.

Wolny [26] presented a modified distributed mobility-adaptive clustering algorithm [27] to adapt the new features of VANET. This modified algorithm is distributed and mobility-adaptive, as well as traffic direction-dependent, and thus, reclustering is avoided when the clusters of vehicles move in different directions. However, this algorithm requires GPS to acquire direction data.

Zhang et al. [11] set packet transfer delay as the relative mobility between multi-hop distanced vehicles and selected vehicles with the smallest aggregate mobility within multi-hop neighbors as CHs. In this scheme, vehicles must identify the aggregate mobility of all N-Hops distance neighbors. Consequently, numerous extra control messages are generated and broadcasted within the network, which eventually reduces the efficiency of cluster formation.

Ucar et al. [12] introduced a vehicular multi-hop algorithm for stable clustering (VMaSC) based on choosing the vehicle with the least mobility. Mobility is calculated with the difference in speed among neighboring vehicles in multi-hop. However, VMaSC requires the support of GPS or similar location services to obtain mobility data.

Hierarchical cluster analysis (HCA) [13] is a fast, randomized clustering algorithm. Instead of attentively selecting the initial CHs and constructing the most stable clusters, HCA attempts to assemble clusters as fast as possible, which leaves cluster optimization in the maintenance phase. The only limiting factor in cluster size is radio propagation, which is avoided by implementing two-hop clustering. HCA is simple and fast, but it does not support optimization related to vehicle movement pattern that can improve cluster stability and CH duration.

Compared with one-hop clusters, multi-hop clusters can extend the range of cluster coverage and gain additional advantages. Hence, multi-hop cluster is a promising direction in VANETs. However, there still remain some open issues to be solved, which can be divided into

two main aspects in general: 1) Multi-hop clustering scheme should reduce the control overhead in the processes of forming and maintaining the cluster for VANETs. 2) More complex routing protocols should be designed deliberately to satisfy the requirement of multi-hop communication among vehicles in different clusters.

3 Preliminaries

3.1 Cluster structure

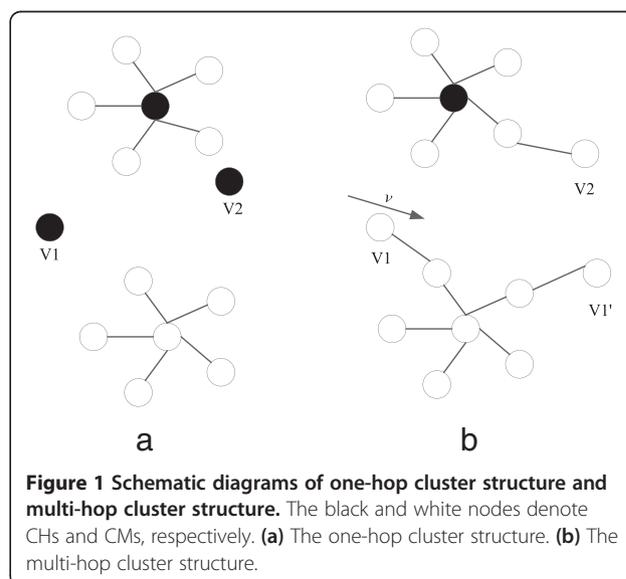
Clusters are virtual groups formed by using clustering algorithms. Each cluster consists of a CH and several CMs. The architecture of a cluster can be divided into two categories, namely, *one-hop* and *multi-hop* clusters, based on routing hops.

In one-hop cluster, each CM can directly communicate with its CH in a one-hop cluster and can directly or indirectly communicate (via CH) with other CMs.

A one-hop cluster structure requires the CH to be directly available to its CMs, which causes many isolated vehicles to emerge after clustering. As shown in Figure 1a, two isolated vehicles (V1 and V2), which are located between two clusters, cannot directly communicate with either CH. Hence, V1 and V2 are divided into two isolated clusters. Isolated clusters increase the number of CHs in a network, which raises communication cost over the network and decreases routing efficiency. In highly dynamic scenarios, a one-hop cluster structure can hardly ensure interconnection between CH and CMs.

In multi-hop cluster, vehicles can communicate with each other in a multi-hop manner.

A multi-hop cluster scheme can decrease the number of CHs to reduce communication cost. As shown in Figure 1b, V1 and V2 can be divided into cluster up



and down, respectively, because a multi-hop structure allows CHs to indirectly communicate with CMs. Moreover, CMs can communicate with their CHs through other intermediate CMs. These conditions allow CMs and CHs to move in a flexible manner.

However, the cost of forming clusters and maintaining a multi-hop cluster structure may be high. A vehicle must determine the aggregate mobility metric of all vehicles in N-Hop distance [11]. This situation generates and broadcasts numerous control messages within the network and reduces the efficiency of cluster formation. Thus, multi-hop clustering algorithms must be designed to improve the efficiency of cluster formation. This objective is realized in this study by presenting a new multi-hop cluster model that depends on local one-hop topology information. The proposed cluster model is presented in the following section.

3.2 Clustering performance metrics

The clustering of MANETs is primarily limited by their energy [22] and processing power. Hence, existing clustering algorithms are optimized for low-resource usage. In VANETs, the scenario is completely different because vehicles are rich in energy resource and highly mobile. Thus, clustering in VANETs mainly aims to improve cluster stability. Improving cluster stability is helpful to reduce the maintenance cost of clusters and increase throughput and routing efficiency.

Cluster stability can be defined through different mechanisms, but the most frequently used are *CH duration*, *CM duration*, *CH change number*, and *cluster number* [10,11].

CH duration is the interval from the period during which a vehicle is selected as a CH to the period it assumes other roles. Similarly, CM duration is the interval between the periods during which a vehicle joins and leaves a cluster. Maximizing the duration of CHs and CMs is useful to improve the stability and minimize the overhead cost of cluster formation [12].

CH change number is the number of vehicles that shift from being CH to other roles. The analysis shows that the cost of selecting CHs and modifying cluster structure is expensive [28]. Consequently, clustering algorithms must attempt to reduce the changing of CHs to decrease the reselection of CHs and the reassociation of CMs. Thus, CH change number is a significant measurable indicator.

The routing efficiency of a VANET exerts the greatest influence. A few clusters accelerate routing in a VANET. Therefore, cluster number is also an important indicator in measuring cluster structure.

3.3 DMCNF algorithm

This section provides a detailed description of DMCNF. First, a novel multi-hop cluster structure model based

on the one-hop neighborhood follow strategy is proposed. A neighborhood follow strategy is then introduced to enable vehicles to independently and adaptively choose which target to follow. Clusters are eventually formed and maintained in a distributed manner based on the strategy. The details of the clustering scheme are also described in this section. Table 1 provides the notations used in this paper.

3.4 Neighborhood follow cluster model

Compared with the traditional cluster structure, a more stable structure is constructed and labeled as the neighborhood follow cluster model, which includes the following properties.

Property 1 Multi-hop. Each cluster comprises a CH and CMs. Each CM connects to its CH directly or indirectly via multi-hop.

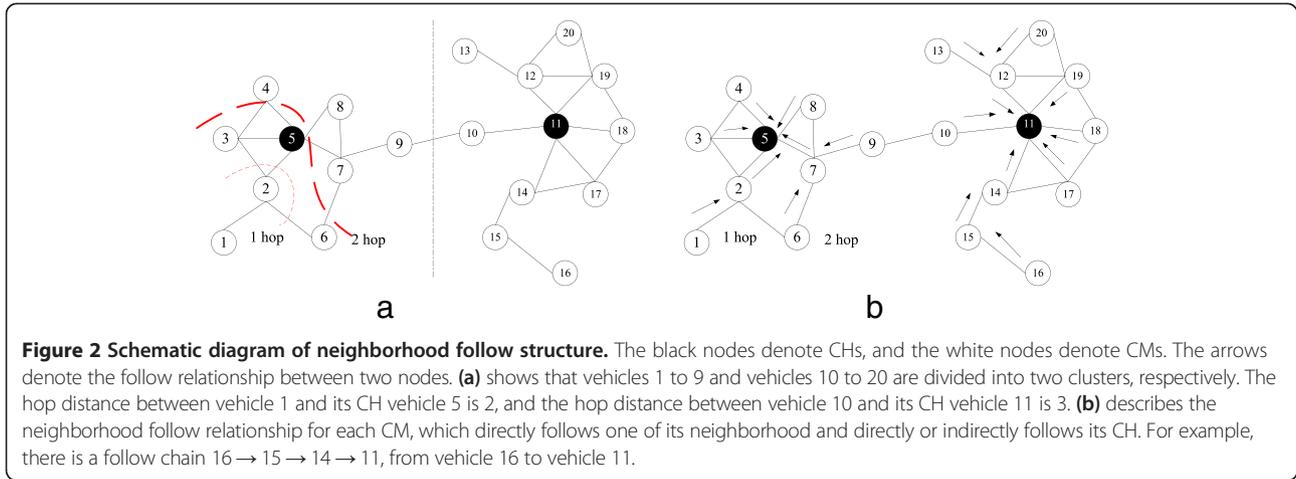
As shown in Figure 2a, the vehicles in a VANET are divided into two clusters, in which vehicles 5 and 11 are the CHs of the two clusters. The CMs for the left cluster include vehicles 1 to 9, with a hop distance of 2 between vehicles 1, 9, and their CH (vehicle 5).

Property 2 A CM passively selects its CH. A following relation exists between two vehicles in a one-hop distance, in which a CM only chooses and follows a stable neighbor and then owns and shares the same CH with its neighbor.

In a traditional multi-hop structure, a CM directly chooses its CH according to the smallest relative mobility when it receives several head messages. This traditional structure requires each CM to be familiar with the relative mobility among all possible CHs. Consequently, extra control messages are widely broadcasted within the network,

Table 1 Notations and description

Notation	Description
$Nblist_x$	One-hop neighbor list of vehicle x
$PktDelay_{xy}$	The packet delay of a packet sent from vehicle y to vehicle x
$RelM_{xy}$	The relative mobility metric between x and y
CH	Cluster head
CM	Cluster member
CH_x	The CH of vehicle x
F_x	The one-hop neighbor follow target of vehicle x
Hello message	The control message sent periodically to maintain the information of neighbors
Hl	The time interval of two continuous periods for a vehicle to send hello messages
Follow message	The notification message sent from a follower to its follow target
Follow reply message	The reply message from a vehicle to its followers caused by a follow message
Head message	The message broadcasted from a new CH



and each vehicle must store and update the relative information of all multi-hop neighbor vehicles. This situation critically increases the overhead.

In large and complex VANETs, a vehicle can hardly obtain the precise details of multi-hop distanced vehicles and decide which CH to choose among multi-hop neighbors. By contrast, a vehicle can quickly obtain the most stable vehicle among its one-hop neighbors. Then, they probably belong to a same cluster. In this study, a vehicle is not required to actively detect its relative mobility with all CHs in multi-hop but must select its CH by following the most stable one-hop neighbor. This mechanism is called the neighborhood follow relationship of belonging to a cluster. Related definitions are provided as follows.

Definition 1 (direct following relationship) Vehicle x exhibits a direct following relationship f , which is defined as follows:

$$f : x \rightarrow y \wedge y \in NBHD(x) \tag{1}$$

where $NBHD(x)$ denotes the one-hop neighbors of vehicle x . A vehicle chooses to follow a target from one-hop neighbors rather than from multi-hop neighbors because obtaining relative mobility with one-hop neighbors is easier and more helpful in reducing extra control messages. As shown in Figure 2, each vehicle selects which target to follow from its one-hop neighbors.

Definition 2 (indirect following relationship) If vehicle y does not belong to $NBHD(x)$, but a following chain exists from x to y , then $x \rightarrow \dots \rightarrow i \rightarrow \dots \rightarrow y$. Accordingly, x indirectly follows y ; such condition is denoted as $x \mapsto y$.

Consequently, a neighborhood follow cluster is defined as follows.

Definition 3 (neighborhood follow cluster)

$$FC_c = \{x | x \rightarrow c \vee x \mapsto c \vee x = c\} \tag{2}$$

where FC_c denotes a cluster whose CH is vehicle c , which is directly or indirectly followed by other vehicles. As shown in Figure 2b, among the CM vehicles in the left cluster, vehicles 2, 3, 4, 7, and 8 directly follow vehicle 5, which is also indirectly followed by vehicles 1, 6, and 9. The proposed cluster structure, which is obtained based on the neighborhood follow relationship, can improve stability in highly dynamic VANETs.

Property 3 Follow uniqueness. A CM directly follows only one one-hop neighbor and directly or indirectly follows only one CH.

Similar to the case of the longest follow chain $16 \rightarrow 15 \rightarrow 14 \rightarrow 11$ in the right cluster (Figure 2), each vehicle in the chain only has one target to follow. That is, no other follow chain exists from vehicles 16 to 11. Moreover, vehicle 16 only indirectly follows vehicle 11.

A cluster structure can be constructed by solving two problems, which are as follows: (1) ‘How does a vehicle decide which vehicle to follow?’ and (2) ‘How are CHs selected according to the neighborhood follow relationship?’ The solutions to these problems are respectively introduced in Sections 4.2 and 4.3

3.5 Neighborhood follow strategy

3.5.1 Relative mobility

Relative mobility is important for a vehicle to decide which target to follow. This study adopts the relative mobility mentioned in [11], which can be calculated based on the beacon (radio propagation) delay on each vehicle.

$$RelM_{x,y} = 10 \log_{10} \frac{PktDelay_{x,y}^{new}}{PktDelay_{x,y}^{old}} \tag{3}$$

where $PktDelay_{x,y}^{old}$ and $PktDelay_{x,y}^{new}$ are the transport delays of two continuous messages and are calculated as follows.

$$PktDelay_{x,y} = Now - TT_y \quad (4)$$

where TT_y denotes the time when vehicle y sends a message to vehicle x .

3.6 Neighborhood follow strategy

Vehicles move fast in a VANET. Consequently, a vehicle may easily change its target if it simply chose it based on instant relative mobility. Thus, when deciding whether it should follow vehicle $PktDelay_{x,y} = Now - TT_y$, vehicle x considers three factors, namely, the relative mobility with y ($RelM_{x,y}$), the current number of followers of y (fc_y), and the historical cluster belonging information of y .

The neighborhood follow strategy is presented after defining two following gains, which are based on the number of followers and the historical cluster belonging information, respectively.

Definition 4 (gain based on the number of followers) A vehicle is inclined to follow a target that is already followed by many vehicles. Thus, when choosing which target to follow, vehicle x considers the numbers of followers of possible targets. Gain is defined by the following expression based on the number of followers.

$$gain_fc_{x,y} = \frac{fc_y}{\sum_{i \in NBHD(x)} fc_i} \quad (5)$$

where $y \in NBHD(x)$. If a vehicle is followed by many vehicles, then these vehicles are relatively stable with it. Consequently, other vehicles are likely to follow vehicles with a stable local topologic structure. It is helpful to improve cluster stability in the dynamic evolution process of a network.

Definition 5 (gain based on the historical cluster belonging information) Vehicle x is likely to follow one target that belongs to the same cluster in which its previous target also belong to. Thus, it is helpful to maintain cluster structure stable. Based on the historical cluster belonging information, gain is defined as follows.

$$gain_his_{x,y} = \frac{\sum_{i \in NBHD(x)} ||RelM_{x,i}|| \cdot \delta(CH_i, CH_y)}{\sum_{i \in NBHD(x)} ||RelM_{x,i}||} \quad (6)$$

$$\delta(a, b) = \begin{cases} 1 & a = b \wedge a \neq \text{null} \\ 0 & a \neq b \end{cases} \quad (7)$$

where CH_x denotes the CH of x whose initial value is null. Formula 6 indicates that gain is greater than 0 when a vehicle and its possible target stay in the same

cluster; otherwise, gain is equal to 0. This condition verifies the aforementioned proposition.

The gain based on the followed number and the historical cluster belonging information is helpful in improving cluster stability.

According to the two gains defined above, the neighborhood follow strategy based on relative mobility and the two kinds of gain is defined as follows.

$$F_x = \arg \min_{y \in NBHD(x)} RelM_{x,y} \cdot (1 - \text{sgn}(RelM_{x,y}) \cdot gain_fc_{x,y}) \times (1 - \text{sgn}(RelM_{x,y}) \cdot gain_his_{x,y}) \quad (8)$$

where F_x denotes the follow target of x . Formula 8 adjusts relative mobility though the two kinds of gains, and then, a vehicle selects its target as the one with the smallest adjusted relative mobility. $RelM_{x,y}$ may be negative, which indicates that x and y are approaching and are likely to be in a same cluster. Thus, the symbol factor $\text{sgn}(RelM_{x,y})$ is used to reduce the value of the adjusted relative mobility.

Cluster structure should be adjusted with network evolution. When forming clusters, CHs remain unidentified. Hence, the neighborhood follow strategy only considers relative mobility and gain based on the number of followers. Meanwhile, when maintaining cluster structure, the neighborhood follow strategy also considers the gain based on the historical cluster belonging information. In such case, a vehicle does not frequently change its target.

3.7 CH decision rule

Direct and indirect neighborhood follow relation between vehicles can be obtained with the strategy. To obtain the final cluster structure, CHs should be selected.

CHs typically forward packets for communication among vehicles. Selecting stable CHs is beneficial to promote routing efficiency and reduce packet loss probability. According to the neighborhood follow strategy, vehicles with more followers and smaller average relative mobility with their neighbors are more stable. Consequently, these vehicles are suitable to be selected as CHs. The rule in selecting a vehicle to be a CH is described as follows.

A vehicle can be considered as a CH according to the following decision rule.

$$bCh_x = \begin{cases} \text{true} & x \rightarrow y \wedge fc_x > fc_y \\ & \wedge AvgRelM_x > AvgRelM_y \\ \text{false} & \text{else} \end{cases} \quad (9)$$

$$AvgRelM_x = \frac{\sum_{y \in NBHD(x)} RelM_{x,y}}{|NBHD(x)|} \quad (10)$$

As indicated in Formula 9, if vehicle x can be considered as a CH, then the number of its followers (fc_x) should be greater than that of its target vehicle y , and its

average relative mobility ($AvgRelM_x$) should be less than that of vehicle y . The basis of this rule is obvious. In a following chain, downstream vehicles are prone to be stable, i.e., they have more followers and are more stable than their followers. However, CHs are exceptions. The follow uniqueness property of the proposed cluster model causes CHs to have their own targets to follow from neighbors in one-hop. Nevertheless, these CHs have evidently greater fc and less $AvgRelM$ than their neighbors.

3.8 Cluster formation and maintenance

This section discusses the detailed processes of forming and maintaining the cluster structure of VANETs based on the neighborhood follow strategy and CH decision rule. DMCNF is a distributed clustering method. Meanwhile, a VANET exhibits a dynamic evolution that presents a dynamic cluster structure, in which each vehicle updates its follow information and dynamically changes its state. Vehicles have three kinds of state, namely, CLUSTER_HEAD, CLUSTER_MEMBER, and CLUSTER_UNDECIDED.

When vehicle x intends to join a VANET, it initializes its state as CLUSTER_UNDECIDED. Then, this vehicle broadcasts a hello message to its one-hop neighbors. Each hello message consists of the id of a vehicle, the id of its CH, the number of its followers, and a forwarding history record list. The element of this list contains the id of a forwarding vehicle and its forwarding time. The time is used to calculate the delay of a packet sent from a vehicle to its one-hop neighbor vehicles. $RelM_{x,y}$, where y is one of the neighbors of vehicle x , can be calculated after a same packet is sent twice between the two vehicles. After receiving a hello message, vehicle x updates its $Nblist$, including the $RelM$ values with its neighbors. The detailed fields of $Nblist$ are shown in Table 2. The process of receiving a hello message is presented in Algorithm 1.

Table 2 Fields of element y of the $Nblist$ of vehicle x

Fields	Description
id	The sequence number of y
$pktDelay$	The packet transmission delay between x and y
$state$	The current state of y
ch	The cluster head of y
fc	The number of followers of y
$relm$	The relative mobility between x and y
$avgRelM$	The average relative mobility between y and each of its neighbors
$followFlag$	FOLLOWER: y follows x FOLLOWEE: x follows y NULL: no follow relation between them

Algorithm 1. Reception of hello message

1. $source_id$: the source id of the packet
2. $reply_id$: the packet is received from vehicle $reply_id$
3. $recv_id$: this vehicle's id
4. $history$: the forward history of the packet
5. $nblist$: the neighbor list of this vehicle
6. $state$: this vehicle's state
7. Get the nearest transmission time tt from vehicle $recv_id$
8. $pktdelay = now - tt$
9. **if** $pktdelay$ appears in $nblist[reply_id]$ **then**
10. Calculate the relative mobility metric and set the value to $nblist[reply_id].relm$
11. **end if**
12. //Whether drop the packet
13. $forward_cnt = 0$
14. **for each** ele in $history$ **do**
15. **if** $ele.id == recv_id$ **then**
16. $forward_cnt ++$
17. **end if**
18. **end for**
19. **if** $forward_cnt == 2$ **then**
20. Drop the packet
21. **else**
22. //Record the forwarding time
23. $history.insert(recv_id, now)$
24. Send the packet to vehicle $reply_id$
25. **end if**
26. **if** $state == CLUSTER_UNDECIDED$ **then**
27. $state = CLUSTER_MEMBER$
28. **end if**

The current state of vehicle x , i.e., CLUSTER_UNDECIDED, transforms into CLUSTER_MEMBER after it sends a hello message. Then, the vehicle starts a timer. After a certain period, referred to as the local learn interval, the vehicle stops receiving hello messages and selects which target to follow. The vehicle chooses a target from the $Nblist$ according to the neighborhood follow strategy and sends a follow message to the target vehicle to notify it about the selection. After receiving the follow message, the target vehicle sends back a follow

Table 3 Simulation parameters

Parameters	Value
Simulation time	400 s
Area range	1,000 m \times 1,000 m
Maximum velocity	10 to 35 m/s
Number of vehicles	100
Transmission range	100 to 300 m
Propagation model	Two-way ground model
Channel	Mac/802_11
LLI	200 ms
CI	5 s

reply message to vehicle x . The follow reply message contains the id , $state$, ch , and fc of the target vehicle. After receiving the follow reply message, vehicle x updates its $Nblist$ with the information carried in the message.

CHs are passively selected in DMCNF. After sending the reply message to the follower, the target vehicle, whose current state is CLUSTER_MEMBER or CLUSTER_HEAD, triggers a process to determine whether to be a CH by adopting the predefined CH decision rule. If the vehicle is determined to be a CH, it changes its state to CLUSTER_HEAD and broadcasts a head message to its followers and the target. Algorithm 2 demonstrates the process of receiving a follow message.

Algorithm 2. Reception of follow message

1. $follower_id$: the follower's id
 2. cur_id : the id of this vehicle
 3. $state$: the current state of this vehicle
 4. $nblist$: the neighbor list of this vehicle
 5. $curAvgRelM$: the average of all the $RelM$ values between this vehicle and its neighborhood
 6. $flwAvgRelM$: the $AvgRelM$ of this vehicle's follow target vehicle
 7. cFc : the number of the followers of this vehicle
 8. fFc : the number of the followers of the following vehicle
 9. //Update the $nblist$
 10. **if** the follower does not appear in the $nblist$ **then**
 11. Add the follower into $nblist$
 12. **end if**
 13. $nblist[follower_id].bFollower = true$
 14. Send a follow reply to the follower
 15. $bCh=false$
 16. **if** $state \neq CLUSTER_UNDECIDED$ **then**
 17. //Then this vehicle might be a cluster head
 18. **if** $cFc > fFc$ and $curAvgRelM > flwAvgRelM$ **then**
 19. $bCh = true$
 20. **end if**
 21. **end if**
 22. **if** $bCh == true$ **then**
 23. Broadcast a head message to followers and the follow target of this vehicle
 24. **end if**
-

Upon receiving a head message, the followers with the state of CLUSTER_MEMBER update their CH information and continue to broadcast the head message to their followers (if any). Some followers may be the CHs of other clusters, and these followers directly drop the head messages. If the target vehicle of the CH is also a CH, then the target vehicle is no longer fit to be a CH. Subsequently, it changes its state to CLUSTER_MEMBER, triggers the process of choosing a target, and directly drops the head message. Algorithm 3 presents the process of receiving a head message.

Algorithm 3. Reception of a head message

1. ch_id : the cluster head's id
 2. ch : the ch id of this vehicle
 3. $state$: the current state of this vehicle
 4. $nblist$: the neighbor list of this vehicle
 5. **for** each v in $nblist$ **do**
 6. **if** $nblist[v].followFlag == FOLLOWER$ and $state == CLUSTER_HEAD$ **then**
 7. $state = CLUSTER_MEMBER$
 8. Boot the follow processing and drop the received packet
 9. **return**
 10. **end if**
 11. **end for**
 12. **if** $state == CLUSTER_MEMBER$ **then**
 13. Drop the packet
 14. **else**
 15. $ch = ch_id$
 16. Broadcast a header message to followers
 17. **end if**
-

Once each vehicle selects a target, its state turns to CLUSTER_HEAD or CLUSTER_MEMBER, which indicates that a transient cluster structure is formed.

Each vehicle triggers the process of choosing a target in every hello interval (HI). The maintenance work is then completed in the time interval during which a vehicle finishes its state updating until the next HI occurs. Each vehicle must execute a check job in every check interval (CI) during maintenance. A CM vehicle is tasked to ensure its connection with its target. If such connection is lost, then a CM triggers the process of choosing a new target. For a CH vehicle, the process of selecting a new target is triggered if other CH vehicles exist in its one-hop neighbor.

4 Simulation results

The algorithms are run on NS2 [29] (release 2.35) platform. VanetMobiSim [30] is used to generate realistic vehicle track data. Several detailed simulation parameters are listed in Table 3. Maximum velocity varies from 10 to 35 m/s, and the transmission range varies from 100 to 300 m. DMCNF is compared with N-Hop [11]. N-Hop is a novel multi-hop clustering scheme for VANETs. N-Hop proposes a mobility metric to evaluate the multi-hop relative mobility between two vehicles. Then, N-Hop uses the relative mobility to decide whether a vehicle can be selected as a CH or not. As DMCNF applies a relative mobility measure similar to that of N-Hop, and both schemes are to discover multi-hop clusters in VANETs, we choose N-Hop for performance comparison. The parameter MAX_HOP of N-Hop is set to 3, which is the maximum hop distance between a CM and a CH. The analysis of the stability performance of

the clustering mechanisms is based on four indicators, namely, CH duration, CM duration, CH change number, and cluster number. Then, the distribution of the number of vehicles on different hops from CHs and the overhead of two algorithms are analyzed.

5 CH duration

In this study, CH duration refers to the interval during which the vehicle is in a CLUSTER_HEAD state until it shifts into a CLUSTER_MEMBER or CLUSTER_UNDECIDED state. The average CH duration is computed by dividing the total CH duration with the total number of state changes from CLUSTER_HEAD to CLUSTER_MEMBER or CLUSTER_UNDECIDED. Figure 3 specifies the average CH duration of the proposed scheme and the N-Hop under different transmission ranges and velocity scenarios.

In particular, Figure 3 indicates that CH duration decreases with increasing vehicle velocity. Increased vehicle velocity makes it difficult for CHs to maintain a relatively stable condition with their neighbor vehicles for a long period. If the CHs do not satisfy the conditions for being a CH, then they are demoted into CMs. Nevertheless, with the increase in vehicle velocity, CH duration is moderately reduced in DMCNF than in N-Hop. Meanwhile, CHs in N-Hop must exhibit the least average relative mobility among their N-Hop neighbors. Such condition increases the difficulty of maintaining stable CHs for a long period under high vehicle velocity. In DMCNF, the stability of CHs only considers the relative mobility with their followers and targets, thereby enhancing the robustness of vehicle velocity.

The transmission range factor influences the stability of CH vehicles, and CH duration increases with increasing transmission range. With a wide transmission range, vehicles do not lose connection with their neighbors.

Similar to the case in DMCNF, the vehicle reboots the process of choosing a new target once it loses its target. This situation may reduce the number of followers of CHs, which may be demoted to CMs, and consequently, increases the number of state changes from CLUSTER_HEAD to CLUSTER_MEMBER. Therefore, increasing transmission range is beneficial to cluster stability.

In DMCNF, a vehicle triggers the process of choosing its target every HI to maintain a dynamic cluster structure for a VANET. Figure 3 reveals that CH duration increases with increasing HI. During each HI, a vehicle checks its connection with its target in every CI. The stability of a vehicle may deteriorate after a while; yet, the vehicle will not reboot the process of choosing a new target as long as it does not lose its connection with its target. However, when HI is too small, the vehicle will reselect a target based on its new local topology structure, which may vary from the old one. Consequently, some CHs lose their followers and become CMs. DMCNF is superior to N-Hop under different scenarios.

5.1 CM duration

CM duration is the interval from the time during which a vehicle joins a specified cluster to the time when it leaves the cluster. As shown in Figure 4, vehicle velocity moderately affects DMCNF in terms of CM duration. In DMCNF, a cluster is developed based on the neighborhood follow relation among vehicles. The cohesiveness of a cluster is strong, i.e., most neighbors of a vehicle are from the same cluster. A vehicle may change its target during its transition. Nevertheless, the new target is prone to have the same CH as the old one, which implies that the vehicle still follows the same CH indirectly. Such situation helps increase CM duration in the cluster. In N-Hop, a vehicle easily

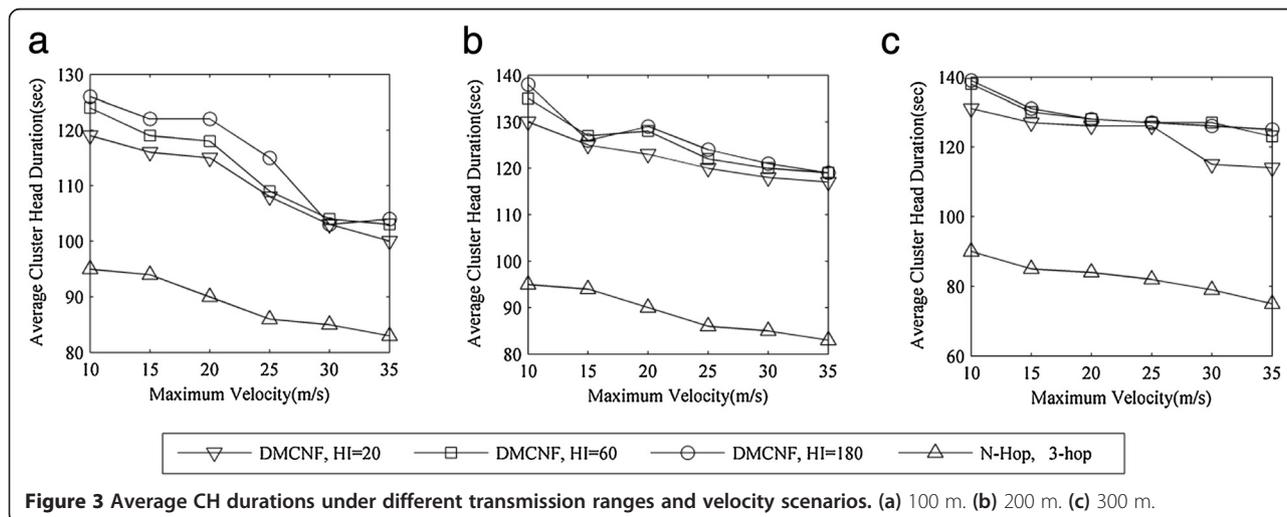
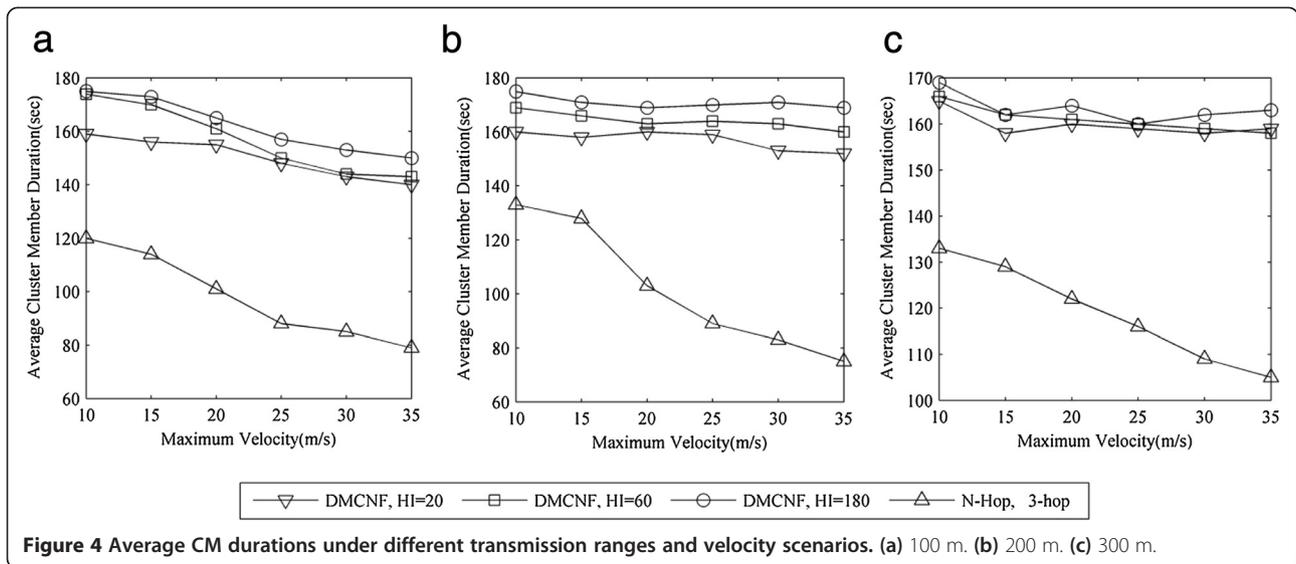


Figure 3 Average CH durations under different transmission ranges and velocity scenarios. (a) 100 m. (b) 200 m. (c) 300 m.



breaks away from its old cluster once it increases relative mobility with its old CH.

CM duration increases with increasing propagation range because vehicles are stable with their neighbors under a high propagation range. Moreover, the increase in HI allows CM vehicles to reduce the frequency of rebooting the process of choosing a target, which allows them to leave their old cluster.

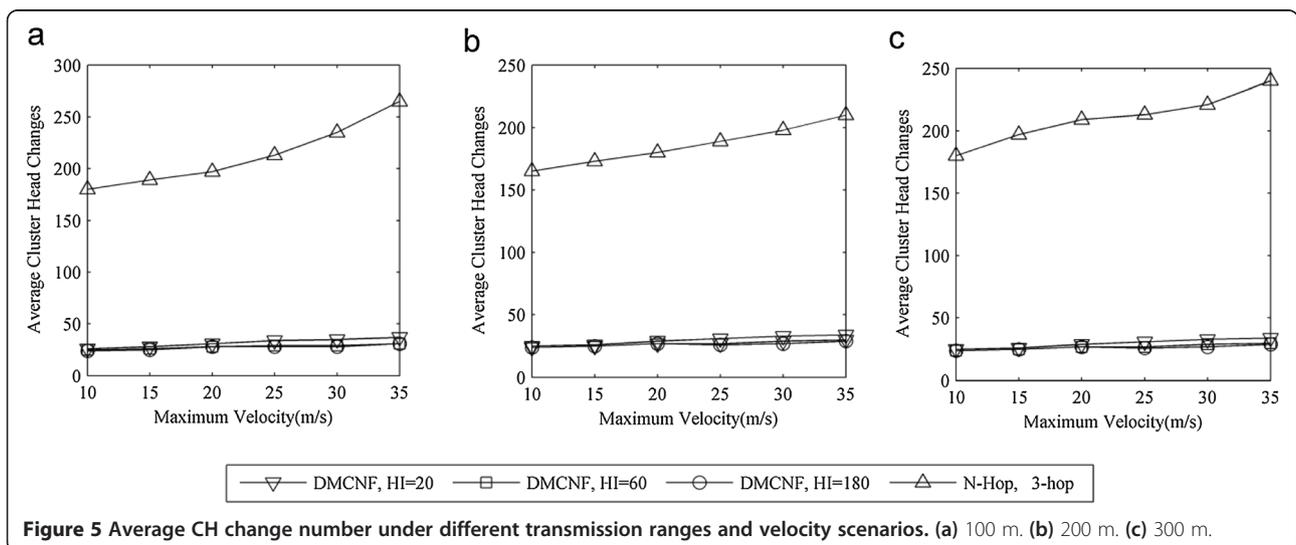
5.2 CH change number

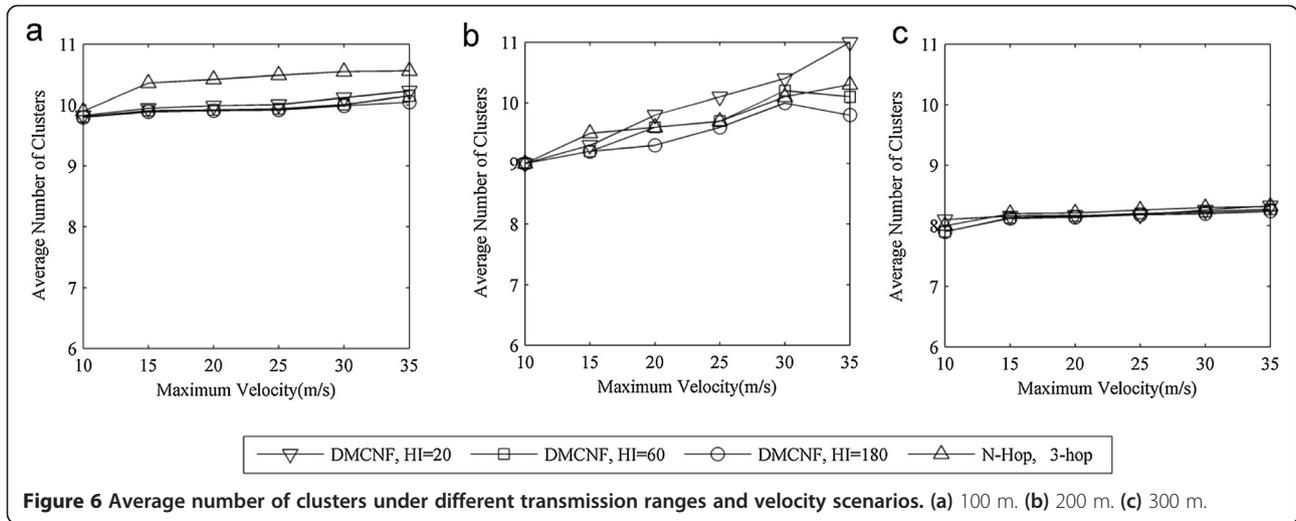
CH change number is the number of vehicles whose state changes from CLUSTER_HEAD to CLUSTER_MEMBER or CLUSTER_UNDECIDED during a simulation process. A low CH change number leads to a stable cluster structure. Figure 5 shows the results of the average CH change number under different transmission

ranges and velocity scenarios. The figure demonstrates that the indicator is thrice more in N-Hop than in DMCNF. Moreover, by using DMCNF, the indicator moderately increases with increasing velocity because a vehicle satisfies the conditions for a CH, which generally has more direct and stable followers (one-hop followers) than CMs. CH is passively reselected when its one-hop neighbors restart the follow process. The target changes of multi-hop neighbors minimally influence CH selection. Consequently, a CH is not easily weakened to other states.

5.3 Number of clusters

A few clusters can improve routing efficiency in a VANET. Figure 6 indicates the average number of clusters under different transmission ranges and velocity scenarios, in





which the average number of clusters is over 50 simulations. The number of clusters in each simulation is the average number of instant clusters that is periodically obtained. As shown in Figure 6, the number of clusters changes minimally with increasing velocity. A vehicle can still identify the cluster to which it belongs to via the multi-hop scheme even if velocity increases and the number of isolated vehicles is reduced. Accordingly, DMCNF can effectively reduce the number of clusters and isolated vehicles, thereby decreasing the number of clusters with increasing propagation range.

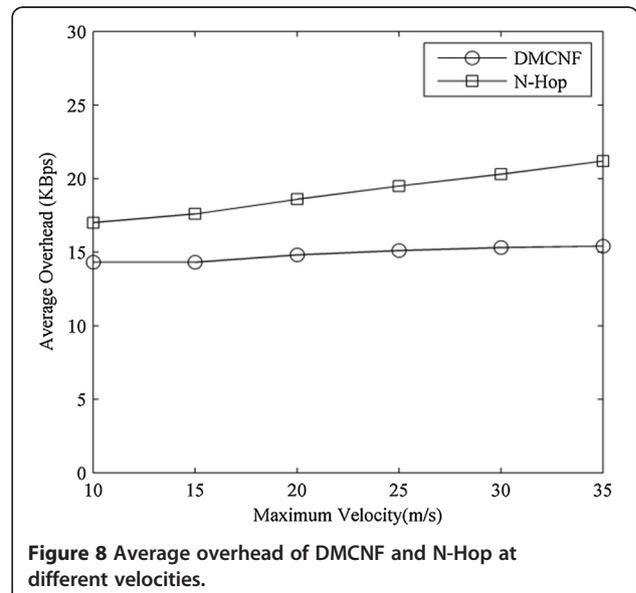
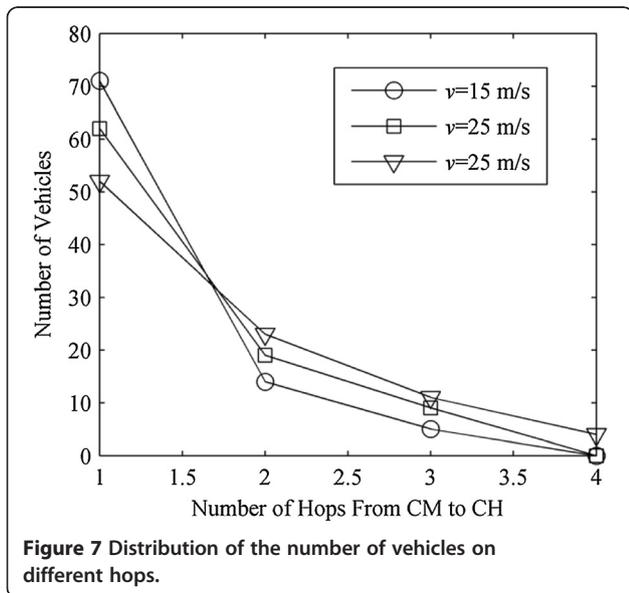
5.4 Analysis of follow situation in cluster structure

In DMCNF, CMs directly or indirectly follow the CH. Clusters are stable if the cohesiveness of the cluster structure is good, and numerous CMs follow the CH

with few hops. Figure 7 illustrates the distribution of vehicles under different hops of the cluster structure obtained by DMCNF. The figure demonstrates that more than half of the vehicles directly follow their CHs and that the number of vehicles decreases with increasing hops, which satisfies the requirement of high cohesion. When the maximum velocity is set to 25 and 15 m/s, the maximum hop is 3. Meanwhile, when the maximum velocity is set to 35 m/s, the relevant maximum hop is 4. These results imply that DMCNF can obtain a cluster structure with high cohesion.

5.5 Analysis of overhead

Figure 8 shows the average control overhead of DMCNF and N-Hop at different velocities. The overhead counts all the control message received by each vehicle in the



entire network and is averaged over the transmission range from 100 to 300 m. Compared with N-Hop, DMCNF performs better in terms of overhead. In DMCNF, each vehicle creates a hello message, a follow message, and a head message during every HI. For the hello message, a vehicle only needs to broadcast a hello message to its one-hop neighbors to calculate the relative mobility between the vehicle and its neighbors. On the contrary, N-Hop needs to broadcast hello messages in multi-hop distance and the size of a hello message increases through the forwarding path, which in turn increases the overhead of VANETs.

6 Conclusions

In this study, a multi-hop clustering scheme with improved stability is obtained. First, the cluster model is presented based on the neighborhood follow strategy. Then, a novel multi-hop clustering algorithm, called DMCNF, is proposed. DMCNF allows vehicles to periodically choose their targets from one-hop neighbors in a distributed manner. The neighborhood follow strategy considers relative mobility, gain based on the number of followers, and gain based on the historical cluster belonging information. Moreover, this strategy improves the stability of clusters during network evolution. The distributed manner of choosing a target leads to easy maintenance of the cluster structure. The efficient routing protocols based on the neighborhood follow cluster structure for VANETs will be explored in the future.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

The authors would like to thank the support of the Technology Innovation Platform Project of Fujian Province under Grant No. 2009 J1007, the Key Project of Fujian Education Committee under Grant No. JK2012003, the Program of National Natural Science Foundation of China under Grant No. 61300104, 61103175, and 61370210, and the Natural Science Foundation of Fujian Province under Grant No. 2013 J01232.

Received: 30 August 2014 Accepted: 10 March 2015

Published online: 01 April 2015

References

- H Hartenstein, K Laberteaux, A tutorial survey on vehicular ad hoc networks. *IEEE Commun Mag* **46**(6), 164–171 (2008)
- S Tayal, M Tripathi, VANET-challenges in selection of vehicular mobility model, in *Proc. of the Second International Conference on Advanced Computing & Communication Technologies* (IEEE, Rohtak-Haryana India, 2012), pp. 231–235
- Z Wang, L Liu, M Zhou, A position-based clustering technique for ad hoc intervehicle communication. *Appl Rev IEEE Transactions Systems, Man, and Cybernetics, Part C* **38**(2), 201–208 (2008)
- RT Goonewardene, F Ali, E Stipidis, Robust mobility adaptive clustering scheme with support for geographic routing for vehicular ad hoc networks. *IET Intell Transp SY* **3**(2), 148–158 (2009)
- E Souza, I Nikolaidis, P Gburzynski, A new aggregate local mobility (ALM) clustering algorithm for VANETs, in *Proc. of Communications (ICC)* (IEEE, Cape Town- South Africa, 2010), pp. 1–5
- A Koulakezian, *ASPIRE: adaptive service provider infrastructure for VANETS* (University of Toronto, 2011)
- Z Rawashdeh, SM Mahmud, A novel algorithm to form stable clusters in vehicular ad hoc networks on highways. *EURASIP J Wirel Commun Netw* **1**, 1–13 (2012)
- S Vodopivec, J Bester, A Kos, A survey on clustering algorithms for vehicular ad-hoc networks, in *Proc. of Telecommunications and Signal Processing (TSP), 2012 35th International Conference on* (IEEE, Brague-Czech Republic, 2012), pp. 52–56
- K Hafeez, L Zhao, Z Liao, B Ma, A fuzzy-logic-based cluster head selection algorithm in VANETs, in *Proc. of Communications (ICC)* (IEEE, Ottawa-Canada, 2012), pp. 203–207. 3–27 11–15 June 2012
- B Hassanabadi, C Shea, L Zhang, S Valaee, Clustering in vehicular ad hoc networks using affinity propagation. *Ad Hoc Netw* **13**, 535–548 (2014)
- Z Zhang, A Boukerche, R Pazzi, A novel multi-hop clustering scheme for vehicular ad-hoc networks, in *Proc. of the 9th ACM international symposium on Mobility management and wireless access* (ACM, Paris-France, 2011), pp. 19–26
- S Ucar, S Ergen, O Ozkasap, Vehicular multi-hop algorithm for stable clustering in vehicular ad hoc networks, in *Wireless Communications and Networking Conference (WCNC)* (IEEE, Shanghai-China, 2013), pp. 2381–2386
- E Dror, C Avin, Z Lotker, Fast randomized algorithm for hierarchical clustering in vehicular ad-hoc networks, in *Ad Hoc Networking Workshop (Med-Hoc-Net), 2011 The 10th IFIP Annual Mediterranean. IEEE*, 2011, pp. 1–8
- C Lin, M Gerla, Adaptive clustering for mobile wireless networks. *IEEE J Sel Areas Commun* **15**(7), 1265–1275 (1997)
- J Yu, P Chong, 3hbc (3-hop between adjacent clusterheads): a novel non-overlapping clustering algorithm for mobile ad hoc networks, in *Proc. of 4th IEEE Pacific Rim Conference on Communications, Computers and signal Processing* (IEEE, Singapore, 2003), pp. 318–321
- P Basu, N Khan, T Little, A mobility based metric for clustering in mobile ad hoc networks, in *Proc. of 21st International Conference on Distributed Computing Systems Workshops* (IEEE, Arizona-USA, 2001), pp. 413–418
- M Ni, Z Zhong, D Zhao, MPBC: a mobility prediction-based clustering scheme for ad hoc networks. *IEEE Trans Veh Technol* **60**(9), 4549–4559 (2011)
- Y Xu, S Bien, Y Mori, J Heidemann, D Estrin, Topology control protocols to conserve energy in wireless ad hoc networks. Center for Embedded Network Sensing, 2003.
- R Selvam, V Palanisamy, Stable and flexible weight based clustering algorithm in mobile ad hoc networks. *Int J Com Sci Inf Technol* **2**(2), 824–828 (2011)
- R Agarwal, D Motwani, *Survey of clustering algorithms for MANET*. *arXiv preprint arXiv:0912.2303*, 2009
- A Bentaleb, A Boubetra, S Harous, Survey of clustering schemes in mobile ad hoc networks. *Commun Netwo* **5**(02), 8 (2013)
- I Humar, X Ge, L Xiang, M Chen, J Zhang, Rethinking energy efficiency models of cellular networks with embodied energy. *IEEE Netw* **25**(2), 40–49 (2011)
- J Volpe, *Vulnerability assessment of the transportation infrastructure relying on the global positioning system*, 2001
- M Thomas, J Norton, A Jones, *Global navigation space systems: reliance and vulnerabilities* (The Royal Academy of, Engineering, 2011)
- S Vodopivec, J Bešter, A Kos, A multihoming clustering algorithm for vehicular ad hoc networks. *Int J Distributed Sensor Networks* **2014**(2014), 8 (2014). Article ID 107085
- G Wolny, Modified DMAC clustering algorithm for VANETs, in *Proc. of 3rd International Conference on Systems and Networks Communications* (IEEE, San Francisco-USA, 2008), pp. 268–273
- S Basagni, Distributed clustering for ad hoc networks, in *Proc. of 4th International Symposium on Parallel Architectures, Algorithms, and Networks (PAAP)*. 1999, pp. 310–315
- M Ni, Z Zhong, K Wu, D Zhao, A new stable clustering scheme for highly mobile ad hoc networks, in *Proc. of 2010 IEEE Wireless Communications and Networking Conference (WCNC 2010, IEEE, Sydney-Australia, 2010)*, pp. 1–6
- K Fall, K Varadhan, The network simulator (ns-2). http://nslam.isi.edu/nslam/index.php/Main_Page
- M Fiore, J Harri, F Filali, C Bonnet, Vehicular mobility simulation for VANETs, in *Proc. of 40th Annual Simulation Symposium (ANSS'07, Norfolk-England, 2007)*, pp. 301–309