



# Stable Clustering Algorithm for Routing Establishment in Vehicular Ad-Hoc Networks

Jieying Zhou, Pengfei He<sup>(✉)</sup>, Yinglin Liu, and Weigang Wu

School of Data and Computer Science, Sun Yat-sen University,  
Guangzhou, China

{isszjy,wuweig}@mail.sysu.edu.cn,  
{hepf3,liuylin6}@mail2.sysu.edu.cn

**Abstract.** With regard to the complex and varied urban scenes in Vehicular Ad-Hoc Network, such as the vehicle nodes with fast speed, unstable links and frequent changes in network topology, this paper proposed a stable clustering algorithm to establish routing for VANET. In this algorithm, clustering was first formed according to the Similar Neighbor Node Table. Then on the basis of the Highest Connectivity Algorithm, parameters such as position and speed of the node were introduced to calculate the Selection Priority which were used to produce the preferred and the alternative cluster head node. The introduction of alternative cluster head nodes improved the stability of clustering to a certain extent. Finally, the clustering was established and maintained for six special scenarios. Compared with the traditional clustering algorithm in VANET, it had the characteristics of lower end-to-end delay, higher packet delivery rate and lower change rate of preferred cluster head nodes, which greatly improved the communication quality of VANET.

**Keywords:** Vehicular Ad-Hoc Network · Routing protocol · Clustering algorithm

## 1 Introduction

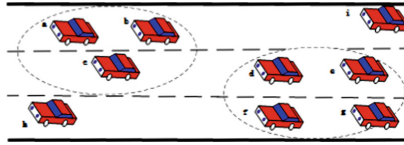
People increasingly place their hope of alleviating or even completely solving traffic problems on intelligent transportation system and vehicle self-organizing network technology [1]. The Vehicular Ad-hoc Networks (VANET) builds a centerless, multi-hop, self-organizing mobile communication network with many nodes [2, 3]. And clustering algorithm is the key content that determines the performance of the VANET [4].

However, at present, there is no clear standard for the routing protocol or clustering algorithm in VANET industry. The clustering algorithm in VANET usually draws lessons from the clustering algorithm in the Mobile Ad-hoc Network (MANET). The representative algorithms are: Minimum ID Algorithm (ID-Lowest) [5], Highest Connectivity Algorithm (HC) [6], WCA [7], AOW [8], ALM [9] and so on. The advantages of ID-Lowest are as follows: the implementation process is simple and convenient; the change rate of cluster head nodes is low. However, its shortcomings are also obvious: because only the minimum ID is considered as the cluster head, it will be

unfair and unstable, and there is no good maintenance mechanism to manage clustering. The HC [6] is similar to ID-Lowest. The disadvantage of HC is that if there is no limit on the number of members in the cluster, it can be a lot. If the number of clusters becomes smaller, then the channel utilization will become very low, the throughput will decrease rapidly.

Although VANET and MANET have some similarities, there are still some differences between VANET in traffic roads. In this paper, according to the characteristics of urban traffic conditions, and its research and analysis, a stable clustering algorithm is designed, which is more suitable for VANET. Compared with HC, ID-Lowest and other clustering algorithms, it has smaller preferred cluster head node change rate and more lasting preferred cluster head node maintenance time, which improves the packet delivery rate, reduces the topology change rate, and can reduce the network transmission delay. The packet delivery rate is improved, which greatly improves the communication quality in the VANET.

## 2 Scenario of the Stable Clustering Algorithm



**Fig. 1.** The cluster of vehicles in local road sections

As shown in Fig. 1, in a road section with three lanes, a cluster was formed because vehicles *a*, *b* and *c* were close, and their average speed was similar. Because the speed of vehicle *h* was fast, even if it was added to the cluster, it would get out of the cluster quickly, so the vehicle *h* was excluded from the cluster. Similarly, vehicles *d*, *e*, *f* and *g* formed a cluster. Because the average speed of *i* was slow, it was also excluded from cluster. Vehicle *i* needed to find other vehicles to form a cluster. Based on this clustering idea, it designs a stable clustering algorithm, which can effectively reduce the network size in the urban scene, reduce the change rate of network topology, and improve the communication quality of the network.

## 3 Mechanism of the Stable Clustering Algorithm

In this stable clustering algorithm, it is assumed that all vehicle nodes in the VANET can get the following information:

- The direction, speed, the longitude and latitude of the location and other information of the vehicle by related devices, such as GPS, on-board sensor devices, etc.
- The HELLO messages used to maintain the topology of the network from the vehicle nodes in their communication range.

- The Beacon message within limited number of hops which includes vehicle ID, position, direction, speed and speed standard deviation, etc.

### 3.1 Clustering for the Similar Neighbor Nodes

In the range of communication between vehicular nodes, the neighbor vehicle nodes obtain each other's basic information through Beacon message. Only the vehicle nodes whose driving speed is within a certain confidence interval are selected to form a similar cluster of neighbor vehicle nodes. The traffic roads in the city will be quite straight and regular. In order to facilitate analysis, only the vehicles driving in the same direction are selected to form a cluster. It is defined that vehicle node  $m$  has the following related parameters:

- $j$ : the number of neighbor vehicle nodes of the vehicle node  $m$ ;
- $\Delta T$ : the sampling period of obtaining the speed of the vehicle node;
- $v_m(t), \sigma_m^v$ : the speed at a certain  $t$  time and the standard deviation of speed;
- $\bar{v}_m$ : the average speed over the periodic sampling interval;
- $\sigma_{max}^v$ : the maximum standard deviation of speed in neighbor nodes;
- $V_{min}^{mj}, V_{max}^{mj}$ : the minimum and maximum driving speed of vehicle nodes;
- $\mu$ : the adjustable coefficient of the size of the speed range.

The vehicle node  $m$  records its speed in the time interval  $\Delta T$ . Calculate the average speed and the standard deviation of the vehicle node  $m$ :

$$\bar{v}_m = \frac{1}{\Delta T} \sum_{t=0}^{\Delta T} v_m(t) \quad (1)$$

$$\sigma_m^v = \sqrt{\frac{1}{\Delta T} \sum_{t=0}^{\Delta T} (v_m(t) - \bar{v}_m)^2} \quad (2)$$

Since it is assumed that the number of neighbor vehicle nodes of the vehicle node  $m$  is  $j$ , the vehicle node  $m$  needs to find a maximum speed standard deviation among the  $(j + 1)$  vehicle nodes. Calculate the maximum speed standard deviation:

$$\sigma_{max}^v = \max\{\sigma_m^v, \sigma_k^v | k = 1, 2, \dots, j\} \quad (3)$$

The confidence interval is used to exclude a few vehicle nodes whose speed is very different from that of most vehicle nodes. On the basis of (3), an adjustable coefficient  $\mu$  is introduced, and the minimum and the maximum speed of vehicle nodes in the confidence interval of Gaussian Mixture Model can be obtained:

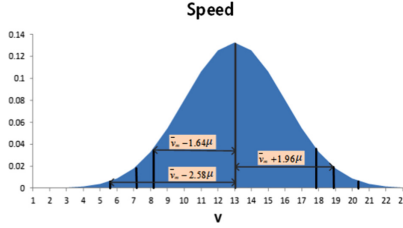
$$V_{min}^{mj} = \bar{v}_m - \mu \sigma_{max}^v \quad (4)$$

$$V_{max}^{mj} = \bar{v}_m + \mu \sigma_{max}^v \quad (5)$$

The value of  $\mu$  can be adjusted according to the average speed of the vehicle and the actual traffic environment. From the properties of the Model, the range of the

confidence interval and the value of the adjustable coefficient  $\mu$  can be obtained (Fig. 2):

$$\begin{cases} P\{|X - \bar{v}_m| < c\} = 2\Phi\left(\frac{c}{\sigma}\right) - 1 \\ P\{|X - \bar{v}_m| < 2.58\mu\} = 2\Phi\left(\frac{2.58\mu}{\sigma}\right) - 1 \approx 99\% \\ P\{|X - \bar{v}_m| < 1.96\mu\} = 2\Phi\left(\frac{1.96\mu}{\sigma}\right) - 1 \approx 95\% \\ P\{|X - \bar{v}_m| < 1.64\mu\} = 2\Phi\left(\frac{1.64\mu}{\sigma}\right) - 1 \approx 90\% \end{cases} \quad (6)$$



**Fig. 2.** The confidence probability of Gaussian Mixture Model

It can be obtained from (6) that some special nodes can be excluded by selecting the appropriate confidence interval. After the confidence interval is set, all the vehicle nodes that satisfy the driving speed in the interval  $V_{min}^{mj} \leq V_j^{mj} \leq V_{max}^{mj}$  are formed into a new set in the neighbor table. Then a similar neighbor node table (SNNT) is created to store and record the information of all vehicle nodes in this collection.

### 3.2 Selection of Preferred Cluster Head Node

The parameters such as the number of neighbor vehicle nodes, speed and position spacing are modeled as a Gaussian Mixture Model. When the number of neighbor vehicle nodes is larger and the driving speed and vehicle position spacing are closer to the average value of the overall vehicle nodes in the table, the greater the selection priority (SP), the greater the probability that the vehicle node will become the preferred cluster head. It is defined that vehicle node  $m$  has the following related parameters:

- $h, v_m$ : the number of vehicle nodes in SNNT and the speed of the vehicle node  $m$ ;
- $v_n$ : the speed of each vehicle node in the SNNT,  $n = 1, 2, \dots, h$ ;
- $\bar{v}_{mean}$ : the average speed of vehicle nodes in SNNT;
- $x_m, y_m$ : the position parameter of the vehicle node  $m$ ;
- $x_n, y_n$ : the location parameters of each vehicle node in SNNT.
- $P_n, \bar{P}_{mean}$ : the Euclidean distance and the average Euclidean distance between  $m$  and the vehicle nodes in the SNNT;
- $\sigma_v, \sigma_p$ : the standard deviation of the speed and the Euclidean distance  $P$  in SNNT;
- $N_v, N_p$ : the normalization of the speed and the Euclidean distance  $P$ ;
- $SP$ : the selection priority of the preferred cluster head node.

Calculate the average speed of vehicle nodes, the Euclidean distance and the average of Euclidean distance between  $m$  and the vehicle nodes in its SNNT:

$$\bar{v}_{mean} = \frac{1}{h} \sum_{n=1}^h v_n \quad (7)$$

$$P_n = \arccos((\sin(Lat_m)\sin(Lat_n)) + (\cos(Lat_m) \cos(Lat_n) \cos(Lng_m - Lng_n))) * R \quad (8)$$

$$\bar{P}_{mean} = \frac{1}{h} \sum_{n=1}^h P_n \quad (9)$$

In (8),  $Lng$  and  $Lat$  represent longitude and latitude, respectively. By using the Z-score standardization method, it is necessary to calculate the standard deviation of speed and Euclidean distance of vehicle nodes in SNNT:

$$\sigma_v = \sqrt{\frac{1}{h} \sum_{n=1}^h (v_n - \bar{v}_{mean})^2} \quad (10)$$

$$\sigma_p = \sqrt{\frac{1}{h} \sum_{n=1}^h (P_n - \bar{P}_{mean})^2} \quad (11)$$

Then, the normalization of speed and Euclidean distance can be calculated:

$$N_v = (v_m - \bar{v}_{mean}) / \sigma_v \quad (12)$$

$$N_p = (P_n - \bar{P}_{mean}) / \sigma_p \quad (13)$$

Finally, the selection priority of  $m$  is calculated:

$$SP = \frac{h}{e^{(|N_v| + |N_p|)}} = \frac{h}{e^{\left( \left| \frac{v_m - \bar{v}_{mean}}{\sigma_v} \right| + \left| \frac{P_n - \bar{P}_{mean}}{\sigma_p} \right| \right)}} \quad (14)$$

### 3.3 Selection of Alternative Cluster Head Node

After the cluster is formed, the most suitable cluster member should be selected as the alternative cluster head node, which makes the cluster structure more stable during the period of maintenance. It is defined that vehicle node  $m$  has the following related parameters:

- $H_n, N$ : the  $N$ th cluster and the number of member nodes in the cluster;
- $\phi_p$ : a set of neighbor vehicle nodes of a member node in a cluster;
- $\mu_i, \theta_i$ : the ID of member node in cluster,  $i$  is a sort index in the range of  $1 - N$ ;
- $C(\mu_i)$ : the coverage cardinality of member nodes in a cluster;
- $\eta(\theta_i), p$ : the  $SP$  of member nodes in a cluster and a member node in the cluster.

The coverage cardinality (CC) represents the number of same nodes in the cluster as the neighbor vehicle nodes of a member node  $p$ :

$$CC_p = \{|H_n \cap \phi_p|, \forall p \in H_n\} \quad (15)$$

Define a set  $\Phi_{I_C}$ , which consists of the  $C(\mu_i)$  of all the member nodes in the cluster  $H_n$ , with a total of  $N$  member nodes:

$$\Phi_{I_C} = \{C(\mu_1), C(\mu_2), C(\mu_3), \dots, C(\mu_N)\} \quad (16)$$

The ID  $\mu_i$  of all the member nodes in the cluster is sorted within the set  $\Phi_{I_C}$ . As  $i$  gets bigger and bigger, the  $CC$   $C(\mu_i)$  are sorted from large to small, corresponding to the ID  $\mu_i$  of the member node. The sort set of the member node ID can be represented as:

$$I_C = \{\mu_1, \mu_2, \mu_3, \dots, \mu_N | C(\mu_k) \geq C(\mu_j), \forall k < j\} \quad (17)$$

The ID  $\theta_i$  of all the member nodes in the cluster are sorted in the set  $\eta(\theta_i)$ . As  $i$  gets bigger and bigger, the  $SP$  of all the member nodes in the cluster  $H_n$  are sorted from large to small. The sort set  $I_P$  of the member node ID can be represented as:

$$I_P = \{\theta_1, \theta_2, \theta_3, \dots, \theta_N | \eta(\theta_k) \geq \eta(\theta_j), \forall k < j\} \quad (18)$$

A new alternative cluster head node has to meet the following two conditions: the vehicle node with a larger  $CC$   $C(\mu_i)$  in (17), that is, the member node in the front of the set; the  $SP$   $\eta(\theta_i)$  of the selected vehicle node is greater than a set threshold  $\eta_T$ . This threshold is selected in the set  $I_P$  of (18) and satisfies the following formula:

$$\eta_T = \eta(\theta_i) \quad i = 1, 2, \dots, N \quad (19)$$

Therefore, a new alternative cluster head node is represented as follows:

$$CH_B = \arg \max\{\mu_i | \eta(\mu_i) \geq \eta_T\} \quad (20)$$

### 3.4 Maintenance of Neighbor Cluster

The maintenance of neighbor cluster is summarized into the following six scenarios, and targeted treatment is made to make the cluster stable.

1. The member nodes in the cluster are separated from the cluster: the ordinary member node in the cluster fails to receive *HELLO* message of the separated node, all the information of the node is deleted from its *SNNT* and the topology structure is updated. The separated node will clear the *SNNT* and change the state to the undetermined state, and regenerate the *Beacon* message to request to join other clusters.

2. The alternative cluster head node is separated from cluster: the member node in the cluster initiates the selection process of the alternative cluster head node and immediately selects a new alternative cluster head node.
3. The preferred cluster head node is separated from cluster: the alternative cluster head node immediately becomes the preferred cluster head node and manages all the members of the cluster. And then re-select an alternative cluster head node.
4. The merging of neighbor clusters: the two preferred cluster head nodes exchange information through *Beacon* messages, and then calculate the degree  $D$  of dissimilarity between the two nodes. When  $D$  is smaller than a certain threshold, the cluster merge is carried out.

$$D = \theta_d \frac{|d_{cb}|}{d_{com}} + \theta_v \frac{|v_{cb}|}{v_c} \quad (21)$$

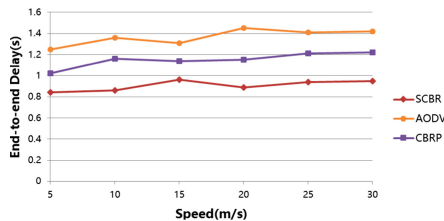
5. Adding the undetermined vehicle nodes to the cluster: when an undetermined vehicle node  $u$  is in the communication range of a cluster,  $u$  will send its *Beacon* message to the preferred cluster head node of the cluster. After receiving the message, the preferred cluster head node  $H$  calculates  $D$ , according to (21), between itself and  $u$ , and then decides whether to accept  $u$  or not.
6. Re-establishment of cluster: the preferred cluster head node needs to obtain the vehicle node information from the *SNNT*, and then calculates its  $SP$  according to the (14). Finally, according to the (22), it is decided whether the vehicle node continues to be the preferred cluster head node:

$$SP = \frac{h}{e^{\left(\left|\frac{v_m - v_{mean}}{\sigma_v}\right| + \left|\frac{P_m - P_{mean}}{\sigma_p}\right|\right)}} < \frac{\pi h}{4e^{\left(\left|\frac{1}{\sigma_v}\right| + \left|\frac{1}{\sigma_p}\right|\right)}} \quad (22)$$

## 4 Simulation and Result Analysis

A routing protocol, the Stable Clustering-Based Routing Protocol in Vehicular Ad-Hoc Networks (SCBR), is proposed based on the stable clustering algorithm which is simulated and tested with the NS2 and the VanetMobiSim.

The driving speeds of vehicle nodes are 5, 10, 15, 20, 25 and 30 m/s, respectively. The number of vehicle nodes is set to 200. According to the simulation results, the influence of different vehicle node speed on the performance index is evaluated.



**Fig. 3.** The effect of speed on end-to-end delay

As shown in Fig. 3, when the number of vehicle nodes is the same, the speed of vehicle nodes has little effect on the end-to-end delay. The main factors affecting the end-to-end delay are the distribution and density of vehicle nodes, the average number of packets forwarding hops and the routing algorithm mechanism of the protocol. When a routing error occurs, the CBRP protocol can be repaired locally, so the end-to-end delay is lower than that of the AODV protocol. The stable clustering algorithm of SCBR protocol can improve the stability of communication link and clustering structure, and ensure that the data packet can reach the destination node successfully. Therefore, using the SCBR protocol, the driving speed of the vehicle node has the least effect on the end-to-end delay.

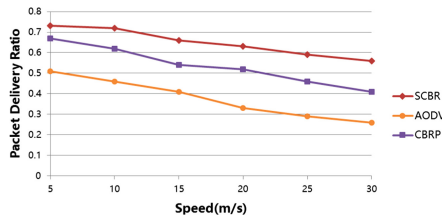


Fig. 4. The effect of speed on packet delivery rate

As shown in Fig. 4, when the vehicle node speed increases, the packet delivery rate of AODV decreases the fastest, while the SCBR decreases more slowly, which is less affected by the vehicle node speed. The main reason is that even when the vehicle nodes are driving at high speed, as long as they can maintain relative stillness with the neighbor vehicle nodes, the cluster structure can be relatively stable and the vehicle nodes can communicate with each other normally. Compared with AODV and CBRP, SCBR has higher packet delivery rate, and its performance index is the best.

## 5 Conclusion

SCBR protocol based on stable clustering algorithm proposed has good performance compared with other relative protocols in urban traffic roads. Clustering algorithm for suburban or high-speed environment could be further studied in the future so that the protocol could be applied to multiple environments simultaneously.

**Acknowledgment.** This work is supported by the National Key R&D Program of China (2018YFB0203803), the National Natural Science Foundation of China (U1801266), and the Program of Science and Technology of Guangdong (2015A010103007).



## References

1. Jin, M.: A brief talk on the development of China's intelligent transportation. *Transp. Sci. Tech.* **20**(2), 140–142 (2013). (in Chinese)
2. Zhang, Y., Wang, H., Peng, L.: Single-layer and cross-layer routing protocols in vehicular ad-hoc networks. *Commun. Technol.* **50**, 2279–2284 (2017)
3. Cheng, J., Ni, W., Wu, W.: A survey of the application research of vehicle ad hoc networks in intelligent transportation. *Comput. Sci.* **36**(1), 1–10 (2014). (in Chinese)
4. Menouar, H., Filali, F., Lenardi, M.: A survey and qualitative analysis of MAC protocols for vehicular ad-hoc networks. *IEEE Wirel. Commun.* **13**(5), 86–94 (2011)
5. Lin, C.R., Gerla, M.: Adaptive clustering for mobile wireless networks. *IEEE J. Sel. Areas Commun.* **15**(7), 1265–1275 (1997)
6. Li, S., Xiao, X.: Vehicle ad-hoc network routing protocol based on road segmentation. *Comput. Eng.* **45**(2), 32–37 (2019)
7. Chatterjee, M., Das, S.K., Turgut, D.: WCA: a weighted clustering algorithm for mobile ad hoc networks. *IEEE J. Cluster. Comput.* **5**(2), 193–204 (2002)
8. Wang, Z.: A vehicle ad-hoc network opportunity routing protocol in 3D urban scene. *Electron Technol.* **47**(7), 60–63 (2018)
9. Togou, M.A., Hafid, A.: Stablecds-based routing protocol for urban vehicular ad hoc networks. *IEEE Trans. Intell. Transp. Syst.* **17**(5), 1298–1307 (2017)