A New Clustering Protocol for Vehicular Networks in ITS

Mengmeng Liu, Xuming Zeng, Mengyao Tao, and Zhuo Cheng(✉)

China University of Geosciences, No. 388 Lumo Road, Wuhan, China
{1227664675,1404347104}@qq.com
{zengxvming,chengzhuo}@cug.edu.cn

Abstract. The dissemination of information is the main application of vehicular networks. The cost of network data transmission is significantly increased and the overall performance of the network routing protocol is deteriorated obviously in the vehicle networks without cluster. In order to ensure the reliability and timeliness of information transmission in VANET, it is necessary to cluster the network in most cases. The existed clustering protocols cannot solve the problem of the stability of the cluster due to the high dynamic mobility in VANET. In this work, we improve a clustering protocol based on the optimization of number of cluster heads, the distance from cluster head to cluster members, and the relative velocity of vehicles within the cluster. The optimized cost function contains three weighting factors for adjusting the specific gravity. The weighting factors can be decided with different demand. Simulation results show that our improved clustering protocol has very good performance for the stability of the cluster. It can provide a good support for the future network routing protocol.

Keywords: Cluster · VANET · Channel allocation

1 Introduction

A vehicular ad hoc network (VANET) is the fusion of the Internet, mobile communication network and the Internet of things, which has a key role in the intelligent transportation systems. In VANET, the idea of clustering is investigated to achieve some of the benefits of an infrastructure-based network without the need for physical infrastructure. Clustering algorithms are proposed to make a hierarchical network structure form in a distributed manner throughout the network. The so-called clustering is to make associating mobile nodes into groups according to some rule set, which is vital for efficient resource consumption and

Supported by CUG2018JM16, Fundamental Research Funds for the Central Universities and SJ-201816, Experimental technology research Funds.

X. Liu et al. (Eds.): ChinaCom 2018, LNICST 262, pp. 615–624, 2019.
https://doi.org/10.1007/978-3-030-06161-6_60

load balancing in large scale networks. Each group is called a cluster which is composed of cluster head (CH) to mediate between the cluster and cluster member (CM). Clustering routing algorithm is based on cluster to achieve node management and data forwarding, which improves routing discovery overhead, broadcast storm and network throughput. The existing clustering algorithms can be classified into four categories: mobility-based, weighted-based, DEA-based or ACS-based clustering, and Location-based clustering protocol.

1.1 Related Work

With the development of the vehicular ad hoc network, the improved protocols from the traditional MANET are difficult to meet VANET?s requirements. There are many routing protocols from different aspects. From the point of clustering protocol, it can be classified into four categories: mobility-based, weighted-based, DEA-based or ACS-based, and Location-based clustering protocol.

(1) Mobility-based clustering protocol: The protocols in this section take relative mobility between a node and its potential cluster heads as cluster head selection parameter, which can be more likely to be able to maintain a stable communication link with small relative velocities when nodes are moving together. Such as [1–4], these algorithms accurately identify nodes showing similar mobility patterns and group them in one cluster.

(2) Weighted-based clustering protocol: In Weighted-based clustering protocol [5–7], cluster head selection for its neighbours is to calculate an index quantifying its fitness. The index, such as the degree of link stability, connectivity, node uptime etc., is a weighted sum of various network metrics. In those protocols, the index representing a node's suitability is based on its entire neighbourhood relationship. When node's connectivity is quite marginal, it may be identified as a best cluster head candidate.

(3) DEA-based or ACS-based clustering protocol: Many DEA-based clustering have been studies in [8–11]. The combination of DEA (Data Envelopment Analysis) and clustering technique is categorized into two methods. The first method is a two-stage algorithm for clustering data by using DEA and a clustering technique such as k-means [9,10]. The second method is to use DEA results for clustering the data [8,11]. ACS-based (Ant Colony System) clustering algorithm is introduced in [12,13].

(4) Location-based clustering protocol: In the designing of routing protocols, considering geographic location information can control the number of cluster heads, reduce the communication cost of inter-cluster and enhance the efficiency of the routing forwarding [14–16]. Reference [15] uses node velocity estimation, introduction of virtual network central node, early warning for cluster head failure and inter-cluster load balancing to form a stable and reasonable clustering structure and achieve the purpose of being easy for frequency planning. Reference [16] is based on Firefly Algorithm and clustering techniques to improve the routing performance under different mobility structure.

1.2 Contribution of This Paper

In this work, a new clustering protocol is proposed to improve the stability and aggregation of network clustering in VANET. Our contribution includes:

(1) The proposed clustering method is related to density. It can be better applied to the different vehicle density environments. The performance of the clustering can well adapt to changes in vehicle density.

(2) The relative velocity of the vehicle is added to the cost function. It can improve the stability of the cluster and provide great stability for network routing protocols over the proposed clustering method.

(3) The weighting factors provide flexible scalability for network routing with different requirements.

2 Network Modeling and Problem Statement

2.1 Network Modeling

In this section, the network model is introduced. The assumptions of considered scenarios are as follows.

(1) We assume that the vehicle is uniformly distributed and its inter-vehicle space is $V_s = N/L$, where N is the total number of nodes and L is the length of highway.

(2) The multi-lane highway scenario is considered as a one-dimensional model.

(3) Each vehicle is equipped with a global positioning system (GPS) to know its own position at any given time.

(4) Each vehicle broadcasts beacon packet every 100 ms which includes its location information.

The straight highway scenario is shown in Fig. 1. There are n vehicles distributed the straight highway with the length of L , denoted as x_i, $i \in \{1, 2, \cdots, n\}$. Each vehicle periodically broadcasts a beacon packet. The probability of p determines that the vehicle can be chosen as a cluster head.

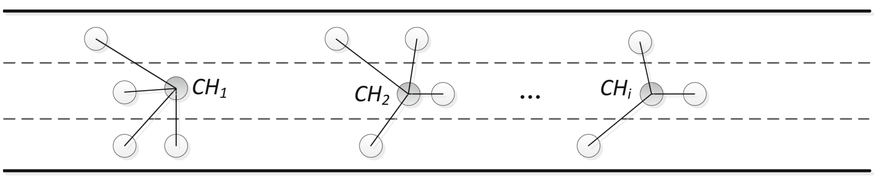

Fig. 1. Demonstration of a clustering scenario.

2.2 Problem Statement

In this work, our problem is to design a clustering protocol that aims to minimize the cost function and lets the nearby nodes divide into the one cluster as much as possible. A better cluster protocol can provide a higher reliability and lower delay for message dissemination.

Let's denote by $C(p)$ the cost function, p is the probability that a node can be chosen as a cluster head. The cost function consists of two parts: the average distance between CH and the distance of a CM to its CH. Therefore, we have:

$$C(p) = e_1 D_{CH}(p) + e_2 D_{CM}(p) + e_3 V(p) \tag{1}$$

$$C = minC(p),\ \ for\ p_{opt} \in (0, 1) \tag{2}$$

where p_{opt} is the optimal value of p, e_1, e_2 and e_3 are the weighting factors.

2.3 Distribution of Relative Velocity

The previous studies show that the velocity of the vehicle traveling on the road is subject to the normal distribution $N(\mu, \sigma^2)$ [17]. Since we only focus on a unidirectional highway scenario, the velocity can be calculated as a scalar. Its probability density function (PDF) can be represented as:

$$f(v) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(v-\mu)^2/2\sigma^2} \tag{3}$$

where v is the vehicle's velocity. In reality, there is a correlation between the velocity of the vehicles. We denote v_i and v_j as any two vehicles' velocities which are jointly normally distributed random variables, then $v_i - v_j$ is still normally distributed. The PDFs of v_i and v_j are $f(v_i) \sim N(\mu_i, \sigma_i^2)$ and $f(v_j) \sim N(\mu_j, \sigma_j^2)$, respectively. So $v_i - v_j$ forms a multivariate normal distribution which can be represented as

$$f(v_{ij}) = \frac{1}{\sqrt{2\pi}\sigma_{ij}} e^{-(v_{ij}-\mu_{ij})^2/2\sigma_{ij}^2} \tag{4}$$

where $v_{ij} = v_i - v_j$, $\mu_{ij} = \mu_i - \mu_j$. Due to the correlation, σ_{ij} can be calculated as

$$\sigma_{ij}^2 = \sigma_i^2 + \sigma_j^2 + 2\rho_{ij}\sigma_i\sigma_j \tag{5}$$

where ρ_{ij} is the correlation coefficient. In particular, whenever $\rho_{ij} < 0$, so the value of σ_{ij}^2 is less then $\sigma_i^2 + \sigma_j^2$.

2.4 Clustering Method

To find the most stable vehicles as the cluster heads (CHs), a clustering method is proposed in this section. We assume that the distance to the source of vehicle x is uniformly distributed over the straight highway scenario with length L [18]. p is the probability that a node will be chosen as a CH. N_{CH} represents the average number of CHs. N_{CM} denotes as the average number of CM that will be merged into a CH. Therefore, we have:

$$N_{CH} = np \tag{6}$$

$$N_{CM} = \frac{n}{np} - 1 = \frac{1}{p} - 1 \tag{7}$$

The requirements for clustering are as follows: (i) to have less number of CHs; (ii) to let each CM within a CH be as close to its CH as possible; (iii) to let the velocities of each cluster?s CMs and CH have less variance. According to all of these requirements, a combined cost function is designed that can measure the cost of the number of CHs, the distance of a CM to its CH, and the variance of relative velocity between a CM to its CH. Therefore, the cost function $C(p)$ is:

$$C(p) = e_1 N_{CH} + e_2 \sum_{k=1}^{N_{CH}} N_{CM} E\{|x_{CH} - x_{CM}|^2\} + e_3 \sum_{k=1}^{N_{CH}} E\{|v_i - \mu|^2\} \tag{8}$$

where e_1, e_2, and e_3 are weighting factors to trade off these three requirements. Due to the uniform distribution of x, we can know that $E\{|x_{CH}|\} = L/2$ and $E\{|x_{CM} - x_{CH}|\} = L/(4np)$. As we know, the variance σ_c^2 of velocity of cluster can be calculated as

$$\sigma_c^2 = E\{|v_i - \mu_c|^2\} \tag{9}$$

where μ_c is the average velocity of all the CMs of a CH in a cluster. Its value can be determined by the speed limit of the highway. Then:

$$C(p) = e_1 np + e_2 np \left(\frac{1}{p} - 1\right) \frac{\Delta d^2}{16 n^2 p^2} + e_3 np \sigma_c^2 \tag{10}$$

where $e_1 + e_2 = 1$, e_3 is selected by the value of σ_c^2. Let $dC(p)/dp = 0$, we find:

$$p^3 + w_1 p = w_0 \tag{11}$$

where

$$w_1 = e_2 L^2 / (16 n^2 (e_1 + e_3 \sigma_c^2)) \tag{12}$$

$$w_0 = 2w1 \tag{13}$$

Note that the polynomial discriminant $D = (w_1/3)^3 + (w_0/2)^2 > 0$, one root is real and the other two are complex conjugates. In terms of the cubic formula, we have:

$$p_{opt} = \frac{w_1}{3\tau} - \tau \tag{14}$$

where

$$\tau = \sqrt[3]{-w_0/2 + \sqrt{w_0^2/4 + w_1^3/27}} \tag{15}$$

It can be verified that $d^2C(p)/d^2p|_{p=p_{opt}} > 0$, which means that $C(p)$ is indeed minimized at p_{opt}, which is the optimal probability for a node to be chosen as a CH.

2.5 Estimation of the Number of Vehicles

The number of vehicles n must be estimated in real time due to the cluster needs to update dynamically with the change of vehicle density. We have assumed that the vehicle is uniformly distributed. So each node can estimate the number of vehicles within the highway in length of L according to transmission range R and the number of vehicles within its range. Let's denote Ψ_t the total number of nodes within a node's transmission range and n_t the total number of nodes within the length of L. We have:

$$n_t = \frac{\Psi_t L}{R}. \tag{16}$$

We can get the value of Ψ_t easily with the help of beacon messages.

Assume that the measurement value of Ψ_t is denoted as $\Psi_t(i)$ during the i-th cluster updating cycle. In terms of Eq. 16, we have:

$$\hat{n}_(i) = \frac{\Psi_t(i)L}{R}. \tag{17}$$

2.6 Clustering Protocol

All vehicles in the target area need to periodically broadcast beacon packets to initialize and update clusters. At the beginning, all the vehicles are in the status: $STANDALONE$. After the establishment of the cluster, the cluster head status is CH, the status of cluster member is CM. The process of clustering is described as below:

(1) All nodes broadcast a beacon packet to the whole nodes within its transmission range, which includes its node ID, GPS coordinates, and direction of travel.

(2) After received the beacon packet, each node stores all the information of the neighbor node and forming a the neighbor information table. Each node estimates the number of nodes $\hat{n}(i)$ in the entire network by using the neighbor information table.

(3) Each node will calculate its own optimized p_{opt} value by Eq. 14 in the updating process of clustering with the estimated value of σ_c^2 and $\hat{n}$. This node generates a random probability p between 0 and 1. If the p_{opt} is greater than the generated random probability p, this node turns to the cluster head.

(4) Each node can decide whether or not to enter the updating process by checking the change of the number of nodes within its neighbor table. If the change value $\hat{n}(i) - \hat{n}(i+1)$is greater than the threshold β_{th}, it will activate the clustering updating process.

(5) If the time since last cluster updating is longer than a predefined constant, activate the cluster updating process.

(6) If a node does not added to any cluster all the time, it remains in the $STANDALONE$ status until the next updating process.

In this protocol, both the number of clusters and the cluster heads are adjusted dynamically with the change of n and σ_c^2.

3 Validation and Simulations

3.1 Simulation Configurations

The simulations are conducted to verify the proposed clustering protocol. The simulator we used is MATLAB version 8.6.

In order to present the performance of the proposed clustering protocol, we design the vehicle that is uniformly distributed in the straight highway with length $L = 2$km. The number of vehicles is $n = 50 \sim 450$. The transmission range is 300m. The range of e_1 is $0 \sim 1$. As a weight factor, it can adjust the weight of the distance between the cluster head and the number of cluster heads. In the analysis of the performance of the clustering protocol, we set up 1000 times runnings to get all of results, and then take its average value. It can obtain the most reliable simulation results. All simulation configuration parameters are shown in the Table 1.

Table 1. Simulation parameters and values

Parameter	Value
Highway length	2 km
Number of lanes	3
Number of vehicles	50∼450
Minimum speed	80 km/h
Maximum speed	120 km/h
Transmission range	500 m

3.2 Simulation Results

In order to show the effect of the clustering algorithm, we use MATLAB to describe the distribution of the vehicle, and then use the proposed clustering protocol to get the CHs. CMs are added to the cluster by the principle of nearest joining. Finally, all the CMs and CHs in the same cluster can be connected directly by lines, so that the readers can directly see the effect of the vehicle network clustering. Figure 2(a) shows the vehicle distribution with no clustering, Fig. 2(b) shows the distribution of vehicles after clustering. As it can be seen from the Fig. 2, our proposed clustering protocol has a very good performance. The adjacent the vehicles are divided into one cluster. It is helpful to reduce the cost of data transmission, including reducing the network delay and improving the reliability of transmission. In order to obtain the performance of the proposed clustering protocol in different network size, the numbers of vehicles we choose are $n = 50, 150, 250, 350, 450$, and the range of weighting factor e_1 is $0.98 \sim 1$. We set $e_3 = 0$ that means the relative velocity of the vehicles is not

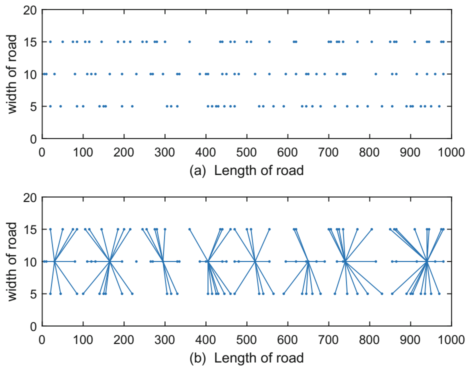

Fig. 2. a Vehicle distribution without clustering. **b** Vehicle distribution with clustering.

considered. After given the values of n and e_1, we can calculate the value of the optimized probability p_{opt} according to Eq. 14. Each node determines whether or not becoming a cluster head by comparing the value of p_{opt} with the random probability. After the CH is determined, other nodes join the cluster by judging the distance from cluster head. The average numbers of CHs in different network size of n are obtained by 1000 simulations. As it can be seen from Fig. 3, when the network size is small, the weighting factor e_1 has small effect on the number of CHs. When the network size is large, the number of CHs is obviously decreasing with increasing of the value e_1. This is in line with the requirements of clustering in the actual vehicle environment, it can not only guarantee the number of clusters in the sparse vehicle network, but also reduce the size of the actual cluster network in dense vehicle network environment. In order to

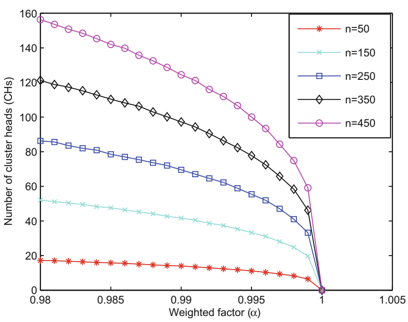

Fig. 3. For the different number of clusters, the values of the weighting factor e_1 have only minor differences when $e_3 = 0$.

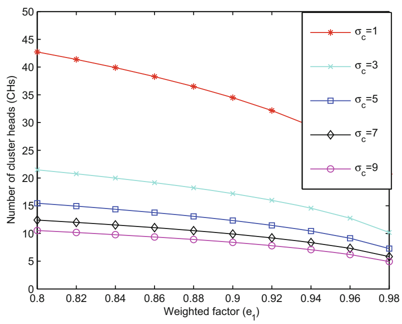

Fig. 4. The number of clusters VS the values of σ_c with $e_3 = 100$.

show the relative velocity of the vehicle on the results of clustering, we choose $\sigma_c = 1, 3, 5, 7, 9$ and $e_3 = 100$. In Fig. 4, as the variance σ_c of the relative velocity of the vehicle within the cluster increases, the number of clusters decreases. In reality, when the relative velocity of the vehicles in the network becomes large, the relative distance of the vehicle will increase. It leads to the requirement for a fast updating of the cluster. By reducing the number of clusters to get a better network topology stability is a good choice.

4 Discussion and Conclusions

In this work, we develop a new clustering protocol for vehicular networks. This clustering protocol is a dynamic clustering protocol. By optimizing the cost function, it obtains the optimal probability of CH. The cost function contains the weight factors, which can be used to adjust the proportion of the two part of the number of cluster heads, the distance of CMs to its CH, and the variance of relative velocity between CMs to its CH. It can achieve the actual needs in different network conditions. The clustering protocol we proposed is adaptive to the different vehicle density and the high dynamic mobility. It ensures that the performance of clustering does not face a significant decline for network routing due to increased vehicle density and the change of velocity.

Our future work would be that we need to find a good method to estimate the number of vehicles in the actual vehicle network and verify the performance and applicability of the clustering protocol by using the actual vehicle traffic data.

References

1. Rawashdeh, Z.Y., Mahmud, S.M.: Toward strongley connected clustering structure in vehicular ad hoc networks. In: 2009 IEEE 70th Vehicle Technology Conference Fall (VTC-Fall), vol. 1–4, pp. 1783–1787. Anchorage, AK, USA (2009)
2. Rawashdeh, Z.Y., Mahmud, S.M.: A novel algorithm to form stable clusters in vehicular ad hoc networks on highways. Eurasip J. Wirel. Commun. Netw. **2012**, 15 (2012). https://doi.org/10.1186/1687-1499-2012-15
3. Daknou, E., Thaalbi, M., Tabbane, N.: Clustering enhancement for VANETs in highway scenarios. In: 2015 5th International Conference on Communications and Networking (COMNET), pp. 1–5 (2015)
4. Kwon, J.H., Kwon, C., Kim, E.J.: Neighbor mobility-based clustering scheme for vehicular ad hoc networks. In: International Conference Platform Technology Service (PlatCon), pp. 31–32. ICT Platform Soc, Jeju, SOUTH KOREA (2015)
5. Daeinabi, A., Rahbar, A.G.P., Khademzadeh, A.: VWCA: an efficient clustering algorithm in vehicular ad hoc networks. J. Netw. Comput. Appl. **34**(1), 207–222 (2011)
6. Vodopivec, S., Beter, J., Kos, A.: A multihoming clustering algorithm for vehicular ad hoc networks. Int. J. Distrib. Sens. Netw. **10**(3), 107085 (2014). https://doi.org/10.1155/2014/107085

7. Hadded, M., Zagrouba, R., Laouiti, A., Muhlethaler, P., Saidane, L.A.: A multi-objective genetic algorithm-based adaptive weighted clustering protocol in VANET. In: 2015 IEEE Congress on Evolutionary Computation (CEC), pp. 994–1002. Sendai, Japan (2015)
8. Thanassoulis, E.: A data envelopment analysis approach to clustering operating units for resource allocation purposes. Omega-Int. J. Manag. Sci. **24**(4), 463–476 (1996)
9. Bojnec, S., Latruffe, L.: Measures of farm business efficiency. Ind. Manag. Data Syst. **108**(1–2), 258–270 (2008)
10. Sharma, M.J., Yu, S.J.: Performance based stratification and clustering for benchmarking of container terminals. Expert. Syst. Appl. **36**(3), 5016–5022 (2009)
11. Po, R.W., Guh, Y.Y., Yang, M.S.: A new clustering approach using data envelopment analysis. Eur. J. Oper. Res. **199**(1), 276–284 (2009)
12. Yang, X.B., Sun, J.G., Huang, D.: A new clustering method based on ant colony algorithm. In: 4th World Congress on Intelligent Control and Automation, pp. 2222–2226., Shanghai, China (2002)
13. Fathian, M., Jafarian-Moghaddam, A.R.: New clustering algorithms for vehicular ad-hoc network in a highway communication environment. Wirel. Netw. **21**(8), 2765–2780 (2015)
14. Edwards, R.M., Edwards, A.: Cluster-based location routing algorithm for inter-vehicle communication. In: IEEE 60th Vehicular Technology Conference, vol. 2, pp. 914–918. Los Angeles, CA (2004)
15. Liu, Z., Ma, Z.X., Shi, R.: Location information based clustering algorithm. Appl. Res. Comput. **28**(12), 4691–4694 (2011)
16. Sachdev, A., Mehta, K., Malik, L.: Design of protocol for cluster based routing in VANET using fire fly algorithm. In: 2016 IEEE International Conference on Engineering and Technology (ICETECH), pp. 490–495. Coimbatore, India (2016)
17. Wang, X., Wang, C., Cui, G., Yang, Q.: Practical link duration prediction model in vehicular ad hoc networks. Int. J. Distrib. Sens. Netw. **11**(3), 216934 (2015). https://doi.org/10.1155/2015/216934
18. Yu, M., David, G.D.: A new traffic aggregation technique based on Markov modulated Poisson processes. In: IEEE Global Telecommunications Conference (GLOBECOM 05), vol. 1–6, pp. 1726–1731. St Louis, MO (2005)