

A Novel Mobility Similarity Measurement Method to Increase the Performance of Community-based Video Delivery in VANETs

SHI-JIE JIA^{1,2}, RUI-LING ZHANG¹, TIAN-YIN WANG¹, LU-JIE ZHONG³
AND MING-CHUAN ZHANG⁴

¹*Academy of Information Technology
Luoyang Normal University
Luoyang, Henan, 471934 P.R. China*

²*State key Laboratory of Networking and Switching Technology
Beijing University of Posts and Telecommunications
Beijing, 100876 P.R. China*

³*Information Engineering College
Capital Normal University
Beijing, 100048 P.R. China*

⁴*Information Engineering College
Henan University of Science and Technology
Henan, 471023 P.R. China*

*E-mail: shjjia@qq.com; ruilingzhang@163.com; wangtianyin79@163.com;
zhonglj@cnu.edu.cn; zhang_mch@haust.edu.cn*

The mobility of mobile nodes is a distinctly important influence factor for video sharing performance, user quality of experience and traffic load remission of core networks in vehicular ad hoc networks (VANETs). In this paper, we propose a novel mobility similarity measurement method to increase performance of community-based video delivery in VANETs (MSMM). In order to accurately represent movement trajectories of vehicles, MSMM calculates relative location between vehicles to refine the geographical location of vehicles. MSMM investigates continuous variation of refined vehicle location to estimate subsection relationship between vehicles and roads and designs a line-segment-based representation method for movement trajectories of vehicles according to the subsection relationship. By building an estimation model of traffic of roads in terms of the hydromechanics and the vehicle following model and by analysis for the historical movement trajectories of vehicles to calculate traffic of roads, MSMM extracts the movement patterns of vehicles. MSMM further respectively designs a recognition method of movement patterns of vehicles and a similarity estimation method of movement behaviors between vehicles, which enables the video requesters to select the video providers which have similar movement behaviors and implement high-efficiency video sharing. We use MSMM to replace the similarity estimation method of node mobility in our previous work PMCV and construct a new video sharing solution (called "M-PMCV") in VANET. Simulation results show how M-PMCV achieves much better performance in comparison with the original solution PMCV.

Keywords: movement behaviors, video sharing, road traffic, VANETs, movement pattern

1. INTRODUCTION

The rapid development of wireless communication technologies (*e.g.* deployed 4G and upcoming 5G) dramatically promotes communication capacities to meet the band-

Received August 29, 2017; revised September 20, 2017; accepted December 12, 2017.
Communicated by George Ghinea.

width demand for the video services in wireless mobile networks [1]. The mobile users can ubiquitously access to the Internet via smart devices (*e.g.*, smart phones and tablet PC) to obtain video content. Fig. 1 illuminates the deployment of video services in vehicular ad hoc networks (VANETs). The video services can provide rich viewing content for the users and attract super-large scale video users. Therefore, the fast increase in scale of video users results in available bandwidth resources in video systems becoming relatively limited for each video user, which brings severely negative influence for system scalability and users' quality of experience (QoE) [2]. In particular, the mobile video users in limited urban area have super large scale and high density and generate huge video traffic [3]. In order to promote system scalability and ensure high user QoE, massive video systems employ mobile peer-to-peer (MP2P) technologies to manage video resources in overlay networks and assign bandwidth resources of peers [4-6]. However, the high cost of managing and fetching videos resources in overlay networks also brings severely negative influences for video sharing performance and user QoE. The virtual community technologies define logical relationship between mobile users and discover the communities composed of mobile users with similar video sharing behaviors in terms of interaction between mobile users [7]. For instance, SMMC in [8] investigates video demand, social relationship and movement behaviors of users to define relationship between users and construct user communities. In fact, the common interests of users for video content can promote successful video lookup probability in communities, which reduces the forwarded number of request messages and decreases delay of video lookup; The similar movement behaviors enable community members which have close geographical distance to achieve high-quality video delivery with low hop count, which reduces transmission delay of video data and relieves traffic load of core networks.

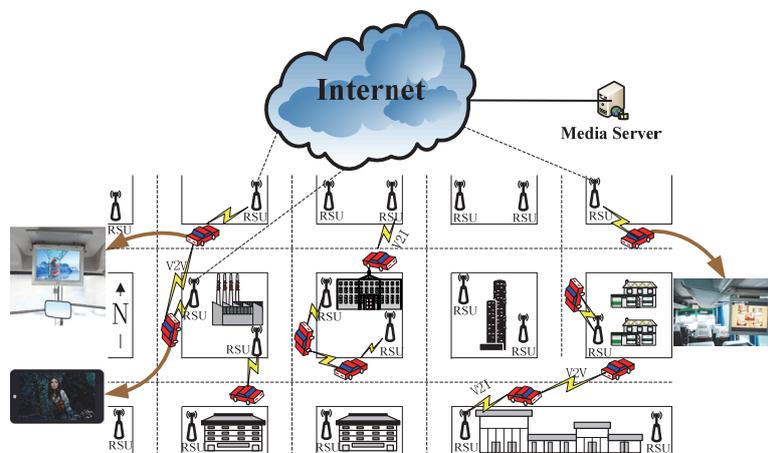


Fig. 1. Video services in vehicular ad-hoc networks.

Obviously, the similarity of movement behaviors between community members is a key factor for promotion of video delivery performance and remission of traffic load of core networks. The accurate estimation for the movement behaviors of mobile users can ensure the high-performance video delivery with low hop count between video re-

questers and providers. In urban area, the passengers in vehicles (*e.g.* personal cars, bus and taxi) have become a kind of video users with huge scale, which watch video content shown by vehicle-mounted video terminals or hand-held smart devices. Unlike the pedestrians, the passengers in vehicles have faster movement speed and higher randomness than the pedestrians, which results in severely negative influence for the video sharing performance and passengers' QoE. There are many studies for the mobility estimation of mobile users in urban area. For instance, PMCV in [9] defines the movement traces of mobile users in terms of the passed access point (AP) and makes use of the Markov process to predict movement trajectories of mobile users in the future. The historical and predicted movement trajectories of mobile users are used to calculate similarity of mobility of mobile users. Cruces *et al.* in [10] designed an estimation method of probability of vehicle encounter, which constructs a contact maps among vehicles to calculate the meeting probability between downloaders and candidate data carriers by analyzing historical data of inter-vehicle communication. Yoon *et al.* designed a trace-driven vehicle mobility models in VANETs, which defines the wireless traces such as WiFi users and APs and makes use of the filtered traces and the built graph to construct the probabilistic mobility model [11]. Kim *et al.* estimated geographical location of mobile nodes in terms of movement traces of mobile devices associating with APs and built the mobility model based on the extracted mobility characteristics of mobile nodes [12]. However, the above methods neglect the influence factors which lead to the abnormal movement behaviors in the stable movement patterns, which results in low estimation accuracy for mobility of mobile nodes. Moreover, the definition methods of vehicular movement trajectories associating with APs do not accurately describe movement trajectories of mobile nodes, which also reduce measurement accuracy of mobility of mobile users. The low-accuracy measurement results for mobility of mobile users not only cause frequent reconstruction of community structure due to the churn of movement state of mobile nodes, but also reduce the video delivery performance between mobile nodes.

In this paper, we propose a **novel Mobility Similarity Measurement Method to increase the performance of community-based video delivery in VANETs (MSMM)**. MSMM designs a refinement method of geographical location of vehicles by calculation of relative location between vehicles. MSMM estimates subsection relationship between vehicles and roads by investigation for the continuous variation of refined vehicle location and formulates a line-segment-based representation method for the movement trajectories of vehicles in terms of the subsection relationship. MSMM constructs an estimation model of traffic of roads in terms of the hydromechanics and the vehicle following model and extracts the movement patterns of vehicles by analysis for the historical movement trajectories of vehicles to calculate traffic of roads. MSMM recognizes the movement patterns of vehicles in terms of current movement trajectories of vehicles and designs a similarity estimation method of movement behaviors between vehicles, which enables the encountered vehicles to find partners and implement video sharing between vehicles with similar movement behaviors. We use MSMM to replace the similarity estimation method of node mobility in our previous work PMCV and construct a new video sharing solution (called "M-PMCV") in VANET. Simulation results show that M-PMCV achieves lower startup delay, lower packet loss rate, high throughput and high video quality than PMCV.

2. RELATED WORK

The community-based video systems group the mobile users with common interests for video content into the same communities in order to promote video sharing performance between nodes in communities. Because the logical link between nodes in communities can be autonomously maintained, the maintenance cost for the whole overlay networks can be distributed into each community, which promotes the system scalability; The nodes in communities make use of the maintained links to fast discover video providers due to common interests between nodes. When the video requesters have found the candidate providers, they need to select the appropriate providers with stable and similar mobility to achieve the low-hop transmission of video data, which reduces data transmission delay and packet loss risk. As Fig. 2 shows, the vehicle A and B are video requester and provider, respectively. When they have similar movement behaviors, the vehicle A can obtain video content with low cost from the vehicle B instead of the server. Therefore, The similarity of mobility of mobile nodes also is considered as a key parameter for the definition of relationship between nodes in communities. The inaccurate estimation for the mobility similarity between mobile nodes results in the fragile relationship between nodes in communities, so that the churn of logical links between nodes in communities causes frequent reconstruction of community structure. This leads to the high maintenance cost for the community structure and reduces the system scalability. Moreover, the inaccurate estimation also reduces video sharing efficiency. For instance, if a video requester connects with a provider which has dissimilar mobility with the requester, data transmission path and geographical distance between them frequently change, which increases packet loss risk and data transmission delay [13].

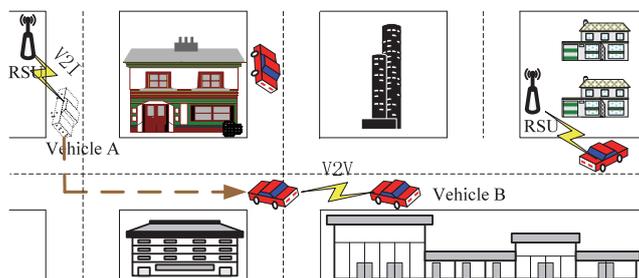


Fig. 2. Video sharing between vehicles with similar mobility.

There are many studies related to the mobility of mobile users in wireless mobile networks. For instance, Lee *et al.* proposed a construction method for the movement patterns of user mobility [14]. The movement trajectories are divided into multiple line segments and the similar line segments are grouped into a cluster by making use of the designed density-based line-segment clustering algorithm. The representative trajectories of clusters are generated in terms of the common characteristics of line segments in clusters, which describes the overall movement of the trajectory partitions in clusters. Yoon *et al.* designed a trace-driven realistic mobility model [11], which uses the association data between WiFi and APs to denote the movement trajectories of mobile nodes. The movement trajectories of mobile nodes can be collected in an actual map and are con-

verted to a graph after the data refinement processing. Based on the built graph, the probability that the mobile users pass the APs in the future is calculated in terms of the extracted movement patterns. Lee *et al.* built a semi-Markov-based mobility model in the temporal and spatial domains in wireless LANs [15]. By making use of semi-Markov model to analyze steady-state and transient behaviors of user mobility, the APs passed by mobile users and the stay time in the future are predicted. PMCV in [9] is a community-based VoD streaming solution over VANETs, which groups the nodes with similar behaviors of playback and movement into the node communities. For the estimation of similarity of movement behaviors, PMCV first makes use of the passed APs in the process of vehicle movement to denote the movement trajectories of vehicles where the time that the vehicles enter the coverage range of APs and the time that the vehicles stay in coverage range of APs are recorded. The movement of any vehicle from an AP to another AP is considered as a transition in the vehicle's state. PMCV employs a Markov process to describe the transition in state of vehicles and calculates the occurrence probability of state transition. The similarity values of residence time in APs between vehicles are calculated in terms of the recorded statistical information. PMCV constructs a matrix of transition probability of vehicle state to predict the movement trajectories of vehicles in the future. The historical and predicted movement trajectories of vehicles are used to estimate the similarity of movement behaviors between vehicles where the similarity values of residence time in APs are used to weight the mobility similarity between vehicles. Cruces *et al.* proposed an estimation method for the possibility of opportunistic encounters of vehicles in urban environment in order to promote data downloading efficiency by leveraging the close geographical distance between the encountered vehicles [10]. The APs passed by the vehicles are considered as the production phase of vehicles and are used to denote the movement trajectories of vehicles. The historical movement trajectories of vehicles are used to calculate the probability that the vehicles go through APs. After the APs exchange the movement information of vehicles such as startup time, speed and direction, they build a contact map to calculate the probability of vehicle encounter. Kim *et al.* built a mobility model from real user traces where the movement trajectories of mobile users is constructed by the APs passed by the mobile users [12]. In order to smooth the movement trajectories of mobile users, the authors employ triangle centroid, time-based centroid and Kalman filter to accurately denote the location of mobile users. Based on the accurate location information of mobile users, the authors extract mobility characteristics in the movement traces of mobile users: pause time, speed, direction, start time, and end time and describe the characteristics of Hotspot regions in urban environment. The mobility model is constructed and is used to calculate the movement destination of mobile users in terms of the probabilities in the region transition matrix.

The most of the above methods for estimation of user mobility employ the definition of movement trajectories associating with the APs, which difficultly ensures the geographical location of mobile users be accurately described. The inaccurate representation of movement behaviors easily results in large errors in the process of extraction of movement patterns of mobile users, so that the accuracy of prediction results for movement behaviors of mobile users in the future difficultly are ensured. Moreover, these proposed methods neglect the random movement behaviors of mobile users. In fact, the movement behaviors of mobile users do not always bring into correspondence with the extracted movement patterns. The random movement behaviors also bring negative in-

fluence for the mobility estimation of mobile users. The measurement for the mobility of mobile users still is challenging.

3. MSMM DETAILED DESIGN

3.1 Representation of Movement Trajectory

The analysis for movement behaviors of mobile users is based on their movement trajectories, so the accurate representation for the movement trajectories can promote precision levels of extraction of mobility characteristics. The most of mobility estimation methods employ the definition of movement trajectories associating with the APs. As Fig. 3 shows, because the RSU has large coverage area, the realistic geographical location of vehicle A difficultly is described by the geographical location of the accessed RSU. The inaccurate representation of geographical location does not accurately describe movement behaviors of mobile users and reflect the intention of drivers in the process of selection of movement paths of mobile users. In fact, the driver in the vehicle A makes the decision for the following travel road in terms of the own travel destination and traffic of roads. The accurate representation of geographical location instead of simple AP-based representation can reflect the intention of mobile users in the process of generation of movement trajectories.

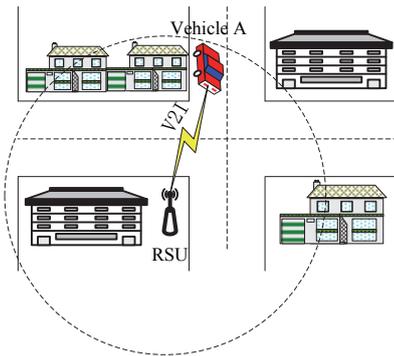


Fig. 3. Representation of movement trajectory associating with Aps.

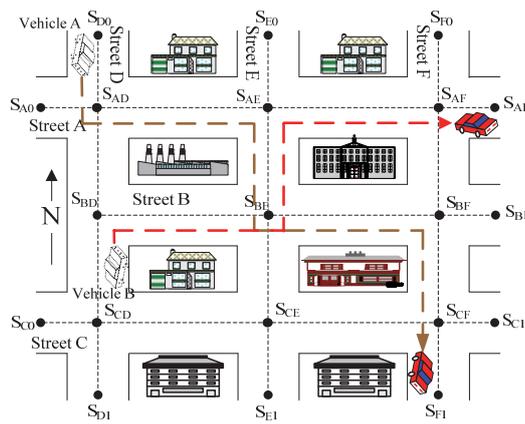


Fig. 4. Representation of movement trajectory associating with line segments.

The vehicles move along the roads in urban environment and the roads are divided into multiple line segments. If the movement trajectories of mobile users are denoted by the line segments associating with the roads, the movement trajectories can be clearly described and the intention of driven route selected by the mobile users can be precisely reflected. As Fig. 4 shows, the roads divide the urban area into multiple the sub-regions and are named by the unique identifications where the horizontal and vertical roads are labeled as “Street A, B, C” and “Street D, E, F”, respectively. If the beginning and end parts of roads are located on boundary of urban area, they are considered as the vertex and have the unique identifications. For instance, the beginning and end parts of street A are named as “ S_{A0} ” and “ S_{A1} ”, respectively. The intersection of the two roads can be

named by the identifications of two roads. The intersection of street A and D is named as S_{AD} . Based on the above naming method of roads, all roads are divided into one or multiple line segments which have the unique identifications. When a vehicle appears on a line segment included in a road, the beginning part of the vehicle is denoted by the line segment in terms of movement direction of the vehicle. For instance, because the vehicle A moves from north to south and locates at the line segment composed of S_{A0} and S_{AD} , the beginning location of the vehicle A is denoted as the line segment $L_{A0 \rightarrow AD}$. Similarly, the beginning location of the vehicle B is denoted as the line segment $L_{CD \rightarrow BD}$ because the vehicle B moves from south to north. There is a key problem of how to obtain accurate geographical location of vehicle A and B. Massive researchers have proposed many solutions related to the vehicle location [16-21]. We employ a simple method to recognize the geographical location of vehicles in terms of the following three conditions.

(1) When a vehicle A has multiple one-hop neighbor nodes in the process of travel, the geographical location of vehicle A can be calculated by the triangle centroid method [12]. If the vehicle A has the multiple one-hop neighbor nodes in the communication range, it exchanges the geographical location information (coordinate) with all neighbor nodes and calculates the geographical distance with the neighbor nodes. The vehicle A makes use of the coordinate values of the two one-hop neighbor nodes (*e.g.* vehicle B and R) which have the closet geographical distance with the vehicle A to refine the own geographical location according to the following equation.

$$\begin{aligned} x_A^* &= \frac{d_{BR}x_A + d_{AR}x_B + d_{AB}x_R}{d_{AB} + d_{AR} + d_{BR}} \\ y_A^* &= \frac{d_{BR}y_A + d_{AR}y_B + d_{AB}y_R}{d_{AB} + d_{AR} + d_{BR}} \end{aligned} \quad (1)$$

Where (x_A, y_A) , (x_B, y_B) and (x_R, y_R) are the coordinate values of vehicle A, B and R, respectively; d_{AB} , d_{AR} and d_{BR} are the Euclidean distance among vehicle A, B and R, respectively; (x_A^*, y_A^*) is the refined coordinate value of vehicle A.

(2) When a vehicle A has only a one-hop neighbor node, the geographical location of vehicle A can be calculated by the line segment centroid method [12]. As the vehicle A only has an one-hop neighbor node (vehicle B) in the communication range, it exchanges the coordinate information with the vehicle B and calculates the centroid coordinate of line segment composed of coordinate of vehicle A and B according to the following equation.

$$x_A^* = \frac{x_A + x_B}{2}, y_A^* = \frac{y_A + y_B}{2} \quad (2)$$

where (x_A, y_A) and (x_B, y_B) are the coordinate values of vehicle A and B, respectively; (x_A^*, y_A^*) is the centroid coordinate of line segment composed of coordinate of vehicle A and B and is considered as the refined coordinate value of vehicle A.

(3) When a vehicle A does not have any one-hop neighbor node, the geographical location of vehicle A is denoted by the GPS coordinate.

Obviously, the location representation based on the GPS coordinate has the lower accuracy than those of methods of triangle and line segment centroid; The triangle centroid method has the higher accuracy than the line segment centroid method. Therefore, the triangle centroid method should be preferentially used in the process of vehicle location awareness in order to avoid the inaccurate description of vehicle location. Moreover, in order to reduce the error of subsection relationship between vehicle location and line segments in roads, the refined location of vehicles can be repeatedly sampled according to the above three conditions. The sampled location of vehicles can be denoted as $L = (x, y, t)$ where x and y are the coordinate values of vehicle location and t is the sampled timestamp. As Fig. 5 shows, the vehicle A has the six location sample data from t_1 to t_6 , which describes the movement trajectories of the vehicle A. The calculated location results of vehicle A do not always locate in the road range (e.g. sample data at t_3 and t_5). Therefore, we firstly define the invalid sample data according to the following rule.

Rule 1: If the vertical dimension from a sample data of vehicle location to the central line in road is less than $3l/2$ where l is the breadth of road, the sample results can be considered as a valid data.

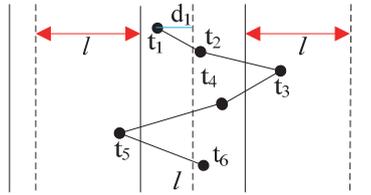


Fig. 5. Multiple samples of vehicle location.

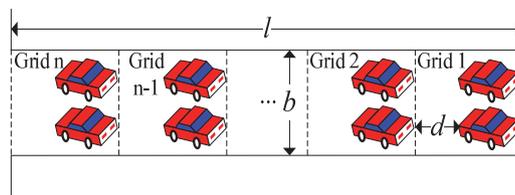


Fig. 6. Grid-based space partition of road.

As Fig. 5 shows, the two sides of road have the two adjacent regions with l breadth, respectively. $3l/2$ is the distance from the central line of road to the boundary of an adjacent region. The definition of adjacent regions is used to promote the fault-tolerant capability for the sample data of vehicle location. The sample location of a vehicle A forms a set $S_{loc} = (l_1, l_2, \dots, l_k)$ and has a time sequence in terms of the timestamp. We employ a simple projection-based movement direction measurement method to estimate conformance levels between vehicle movement direction and road. Firstly, l_1 and l_2 form a line segment and the projection length of the line segment to the central line of road is defined as $PL_{12} = |y_2 - y_1|$ (e.g. the purple line segment in Fig. 5). Similarly, the projection length of line segments composed of l_1 and other items in S_{loc} to the central line of road can be obtained and the projection length set is defined as $S_{PL} = (PL_{12}, PL_{13}, \dots, PL_{1k})$. If the items in S_{PL} meet the linear increasing trend and all sample location is valid, the vehicles move along the road and the movement trajectories of vehicles can be denoted by the line segment of road. Therefore, the whole movement trajectories of vehicles are considered as the set of line segments in roads in terms of the above method. For instance, the whole movement trajectories of vehicle A can be defined as $tr_A = (L_{D0 \rightarrow AD}, L_{AD \rightarrow AE}, L_{AE \rightarrow BE}, L_{BE \rightarrow BF}, L_{EF \rightarrow CF}, L_{CF \rightarrow F1})$, as shown in Fig. 4.

3.2 Estimation of Road Traffic

There are multiple available travel paths from beginning point to ending point in urban environment. When drivers determine the driving destination, they formulate the unique travel plan from multiple candidate plans. The travel cost is very important influence factor for the formulation of travel plan and includes time and journey cost. For instance, as Fig. 4 shows, S_{CD} and S_{A1} are beginning and ending points of vehicle B, respectively. The vehicle B locates at the midpoint in $L_{CD \rightarrow BD}$ and travels from south to north. Therefore, the number of the shortest travel paths from S_{CD} to S_{A1} is 3 and each shortest travel path includes the five line segments. If the vehicle B drives along any shortest path, it can achieve the minimum journey cost. If the traffic congestion is neglected and the three shortest paths have the same travel speed, the vehicle B which drives along any shortest path can obtain the minimum time cost. However, the maximum travel speed allowed by all roads is different and all roads have different capacities to return to normal traffic from traffic congestion. In fact, the traffic of roads is the determining factor for travel speed and resilient capacities of roads. The vehicles unidirectionally travel along the roads, namely the vehicles cannot optionally change the travel direction. Therefore, the roads are considered as the closed conduits. In terms of traffic definition of closed conduits in the hydromechanics, the traffic of any road L_i can be defined as:

$$TR_i = N_{cs} \times v_{cs}, v_{cs} = \frac{\sum_{c=1}^{N_{cs}} v_c}{N_{cs}} \quad (3)$$

where N_{cs} and v_{cs} denote number and average speed of vehicles which pull out cross section of a road L_i in unit time. The travel speed of all vehicles is in $(0, v_{\max}]$ and v_{\max} is the maximum travel speed allowed by the road L_i . The roads with high traffic support high travel speed of vehicles, which reduces the time cost of vehicles; On the other hand, the roads with high traffic allow more vehicles travel, which reduces the risk of road congestion and promotes the resilient capacities of roads. In other words, the roads with high traffic enable the more vehicles travel with high and stable speed. Therefore, the travel plan of vehicles should include as many roads with high traffic as possible. As Fig. 6 shows, a road is considered as a closed conduit and is divided into n grids with the same length and width, where the vehicles occupy much space in each grid and the distance between front and later vehicles is d . N_{cs} can be defined as:

$$N_{cs} = k_{cs} \times \rho \quad (4)$$

where k_{cs} is number of grids which pull out cross section of L_i in unit time and ρ is number of vehicles included by a grid and is a constant. Because the volume of roads is limited, the number of vehicles accommodated by the roads is limited. The volume of a road L_i can be defined as:

$$V_i = N_i \times S_v + S_d \quad (5)$$

where N_i is the number of vehicles on the road L_i ; S_v is the space occupied by a vehicle;

$N_i \times S_v$ denotes the space occupied by all vehicles. $S_d = n \times d \times b$ denotes the space occupied by spacing distance of vehicles. The number of vehicles accommodated by the roads is related with the travel speed of vehicles. Firstly, in terms of the standard of braking distance of vehicles, d can be defined as:

$$d = c_1 \times v^2 + c_2 \times t \times v \quad (6)$$

where c_1 is the braking coefficient of vehicles and also is constant; c_2 is a constant, t is the response time of drivers and v is the travel speed of vehicles. If $v=0$, the value of d also is 0 and $S_d=0$. At the moment, the number of vehicles accommodated by L_i can reach the maximum value N_{\max} . When the vehicles travel with the maximum speed allowed by L_i ($v = v_{\max}$), d reaches the maximum value. However, N_i reaches a critical value $N_c \neq 0$. If $N_i \leq N_c$, $N_i \times S_v + n \times d \times b \leq V_i$, where $N_i \in (0, N_c]$ and $v \in (0, v_{\max}]$. In other words, when $N_i \in (0, N_c]$, the variation of travel speed of vehicles is unrelated with the variation of number of vehicles on L_i (namely all vehicles on L_i can travel with the maximum speed v_{\max}). Therefore, the traffic of L_i is related with the number of vehicles which drive into L_i in unit time. Let v_e and N_e denote the speed and number of vehicles which drive into L_i in unit time, respectively, where $v_e = v_{\max}$. The traffic of L_i can be defined as:

$$TR_i = N_e \times v_{\max}, N_i \in (0, N_c] \quad (7)$$

where v_{\max} is a constant. On the other hand, if $N_i \in (N_c, N_{\max}]$, there is the interactive influence between number and speed of vehicles on L_i . In terms of mass conservation law of the hydromechanics, the increment of traffic is the difference value between traffic values of entering and leaving L_i , namely $N_e \times v_e - N_{cs} \times v_{cs} = \Delta TR_i$. If $N_e = N_{cs}$ and $v_e = v_{cs}$, $\Delta TR_i = 0$. As we know, the vehicles pass through roads following the queueing process, so the front vehicles have the conditionality for the latter vehicles. The speed of latter vehicles always keeps the same change in terms of variation of speed of front vehicles. The conditionality between vehicles also constantly transmit to the subsequent vehicles. When the speed of vehicles which leave roads decreases, the number of vehicles which leave roads in unit time also reduces. If the values of N_e and v_e are constant and $N_e > N_{cs}$, $v_e > v_{cs}$, $\Delta TR_i > 0$, namely the number of vehicles on L_i increases. Otherwise, the number of vehicles which leave roads in unit time also increases with increasing speed of vehicles. If the values of N_e and v_e are constant and $N_e < N_{cs}$, $v_e < v_{cs}$, $\Delta TR_i < 0$, namely the number of vehicles on L_i decreases.

We discuss the representation of variation of traffic in terms of the conductivity of variation of vehicles' speed. In fact, the vehicles on roads can be considered as continuous and homogeneous medium. When the front vehicles which locate at the road exit change their speed (acceleration or deceleration) from v_e to v_{cs} during a time period Δt , the conductivity enables the sequent vehicles make the same change in the speed during Δt . The traffic increment generated by the vehicles on roads can be defined as $\Delta N \times v_{cs}$ where $\Delta N = N_e - N_{cs}$ denotes the increment of number of vehicles leaving L_i during Δt . Moreover, the vehicles which drive into L_i during Δt also change their speed from v_e to v_{cs} . The traffic increment generated by the new vehicles on roads can be defined as $N_e \times \Delta v$ where $\Delta v = v_e - v_{cs}$. Therefore, the increment of traffic of L_i can be defined as:

$$\Delta TR_{L_i} = \Delta N \times v_{cs} + N_e \times \Delta v \tag{8}$$

where $\Delta N + N_e \leq N_{max}$. If $v_{cs} = v_e$ results in $\Delta v = 0$, $\Delta TR_i = \Delta N \times v_e$ denotes the traffic increment generated by the all vehicles which change current travel speed. When the speed of front vehicles reaches the same level with the new vehicles which drive into L_i , the new vehicles do not bring the increment of traffic. If $v_{cs} = 0$ results in $\Delta v = v_e$, $\Delta TR_i = N_e \times v_e$ denotes the traffic increment generated by the new vehicles which drive into L_i . When the speed of front vehicles is 0, all sequent vehicles do not bring the increment of traffic. However, the vehicles have a delay time to make the change in keeping with speed variation of front vehicles in the process of real travel. The speed of sequent vehicles does not reach to v_{cs} during Δt , namely $\Delta v \neq v_e - v_{cs}$ and $\Delta N \neq N_e - N_{cs}$. In terms of vehicle following model [22], the variation of speed of sequent vehicles can be defined as:

$$\Delta v = \frac{(1 + T \times \eta)}{T} (v_{cs} - v_e) \tag{9}$$

where T and η are relaxation time and response coefficient of the latter vehicles, respectively. ΔN can be defined as:

$$\Delta N = \rho / T \tag{10}$$

where $1/T$ denotes the number of grids whose speed happens to change in unit time. Therefore, the increment of traffic of L_i can be re-defined as:

$$\Delta TR_i = \Delta N \times \Delta v + N_e \times \Delta v = \left(\frac{1}{T} \rho + N_e\right) \times \Delta v. \tag{11}$$

By the combination of Eqs. (7) and (12), the traffic of L_i can be defined as:

$$TR_i = \begin{cases} N_e v_{max}, N_i \in (0, N_c] \\ N_e (v_e - \Delta v) - \frac{\rho}{T} \Delta v, N_i \in (N_c, N_{max}] \end{cases} \tag{12}$$

If the vehicles drive into L_i with a constant speed and follow the Poisson distribution, v_e is a constant and $N_e = \lambda$. If the difference of capacity between drivers is neglected, T and η can be set a constant. After Eq. (9) is substituted into Eq. (12), the traffic of L_i is related with variation of v_{cs} . By analysis for the historical travel records of vehicles on L_i , the average values of speed of vehicles which drive into and pull out L_i are considered as the values of v_e and v_{cs} . TR_i can have numerical solution during different time interval. The travel path which has the largest traffic among all shortest paths corresponding to starting and destination points can be considered as the movement pattern of vehicles. Let $MPS = ((S_a, D_a, TS_a, TE_a, P_a), (S_b, D_b, TS_b, TE_b, P_b), \dots, (S_m, D_m, TS_m, TE_m, P_m))$ denote the movement pattern set of vehicles corresponding to combination results of all starting and destination points during different time interval. S_a and D_a denote starting and destination points of vehicles, respectively; TS_a and TE_a denote starting and destination time, respectively; P_a is movement pattern of vehicles.

3.3 Recognition of Movement Patterns of Vehicles

The similarity estimation of movement behaviors between vehicles is the key factor for node community construction and video provider discovery. If the movement behaviors of vehicles belong to the same movement patterns in the process of vehicle movement, the movement behaviors of vehicles can be consistent with high probability. The accurate recognition for movement patterns of vehicles can efficiently reduce error for the consistency levels between movement behaviors of vehicles. There are the following two context for recognition of movement patterns of vehicles.

(1) When the beginning and ending points of vehicles can be known in advance (*e.g.* navigation), the movement patterns of vehicles easily are obtained by making use of the beginning and ending points to search the corresponding items in *MPS*. For instance, the two vehicle A and B have corresponding movement patterns tr_A and tr_B , respectively. The similarity of movement behaviors between them can be obtained according to the following equation.

$$S_{AB} = \frac{\sum_{i=1}^{|tr_A \cap tr_B| - s} D(L_i)}{\text{MAX}[\sum_{e=1}^{|tr_A| - a} D(L_e), \sum_{c=1}^{|tr_B| - b} D(L_c)]} \quad (13)$$

where $L_i \in tr_A \cap tr_B$; $|tr_A \cap tr_B|$ returns number of intersection between tr_A and tr_B ; $D(L_i)$ returns the distance of line segment L_i ; s is the maximum values of number of line segments in $tr_A \cap tr_B$ which have been passed by vehicle A or B; $|tr_A \cap tr_B| - s$ denotes the surplus number of common line segments between vehicle A and B; a and b are the number of line segments which have been passed by vehicle A and B, respectively; S_{AB} is the ratio between distance of common line segments of tr_A and tr_B and maximum distance of line segments of tr_A or tr_B . If S_{AB} is equal or greater than the threshold value V , the vehicle A and B have similar movement behaviors; Otherwise, $S_{AB} < V$ indicates that the movement behaviors of vehicle A and B are dissimilar.

(2) If the beginning and ending points of vehicles cannot be known in advance (*e.g.* privacy protection), the recognition of movement patterns only relies on the prediction in terms of current movement trajectories of vehicles. For instance, as Fig. 4 shows, the vehicle A has moved from S_{D0} to S_{AE} at t_1 . The movement pattern of vehicle A is limited to a subset $SMPS_A$ of *MPS* in terms of t_1 and existing trajectory $tr_A = (L_{D0 \rightarrow AD}, L_{AD \rightarrow AE})$. However, it is very difficult to only make use of t_1 and tr_A to accurately obtain movement pattern of vehicle A. For instance, when the vehicle A moves to S_{AE} , it may have many destination points (*e.g.* S_{E0}, S_{F0}, S_{A1}). Therefore, the travel path which has the largest traffic among all movement patterns corresponding to combination results of the beginning point S_{D0} and all destination points is considered as the movement pattern of vehicle A. For instance, the movement pattern of vehicle A can be defined as $tr_{D0 \rightarrow F1}$. If the number of vehicles which move from multiple beginning points to S_{AE} at t_1 is k , the movement patterns of these vehicles can also be defined in terms of the above method.

When the vehicle A moves to S_{AE} at t_1 , the similarity values of movement behaviors between vehicle A and the k vehicles can be calculated according to Eq. (13). If the similarity value S_{AC} between vehicle A and a vehicle C is largest among the k vehicles and $S_{AC} > V$, the vehicle A and C have similar movement behaviors; Otherwise, if $S_{AC} < V$, the movement behaviors of vehicle A and C are dissimilar. When the vehicle A has the similar movement behaviors with the vehicle C, the vehicle A can decide whether joining into a community including vehicle C or fetching a video from vehicle C. If the vehicle A finds that the vehicle C has left the own one-hop communication range in the process of movement, it re-discovers the new vehicles with similar movement behaviors.

4. TESTING AND TEST RESULTS ANALYSIS

4.1 Testing Topology and Scenarios

We use the proposed estimation method MSMM of mobility similarity to replace the estimation method of movement similarity in our previous work PMCV [9] (The new video sharing solution is called “M-PMCV”). We compare the video sharing performance of M-PMCV with the original PMCV. The two solutions M-PMCV and PMCV are deployed in a VANET-based environment by making use of the NS2. The simulation time is 500 s. We still employ the original settings of main simulation parameters in PMCV and the detailed settings are listed in Table 1. There are 300 mobile nodes in a $1000\ m \times 1000\ m$ square area. The urban topology includes 4 horizontal and 4 vertical streets where every street has two lanes in each direction and the 12 APs equipped with IEEE 802.11p WAVE interfaces are deployed in VANET, as Fig. 7 shows. The four roads: $L_{A0 \rightarrow A1}$, $L_{D0 \rightarrow D1}$, $L_{E0 \rightarrow E1}$ and $L_{H0 \rightarrow H1}$ allow that the maximum values of travel speed of vehicles are 16 m/s and the braking length of vehicles on the four roads is 25 m; The four roads: $L_{B0 \rightarrow B1}$, $L_{C0 \rightarrow C1}$, $L_{F0 \rightarrow F1}$ and $L_{G0 \rightarrow G1}$ allow that the maximum values of travel speed of vehicles are 12 m/s and the braking length of vehicles on the four roads is 15 m. The cross section of each road includes the two vehicles in terms of different travel direction, namely $\rho=2$.

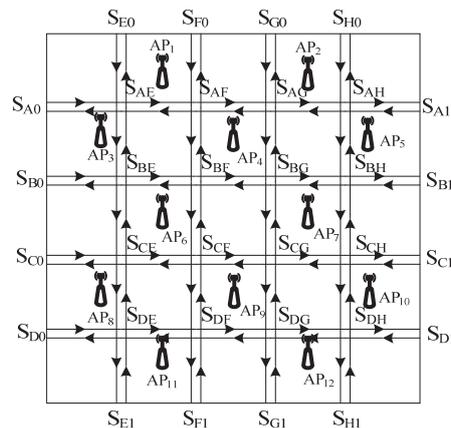


Fig. 7. AP architecture in VANET.

Table 1. Simulation Parameter Setting For VANET.

Parameters		Values
VANET	Area	1000×1000 (m^2)
	Channel	Channel/Wireless Channel
	Network Interface	Phy/WirelessPhyExt
	MAC Interface	Mac/802_11Ext
	Bandwidth	25 Mbps
	Frequency	9.14 GHz
	Multiple Access	OFDM
	Transmission Power	33 dBm
	Wireless Transmission Range	250 m
	Interface Queue Type	Queue/DropTail/PriQueue
	Interface Queue Length	50 packets
	Antenna Type	Antenna/OmniAntenna
	Routing Protocol	DSR
	Mobile Speed Range	[0,20] m/s
	Network Layer Protocol	IP protocol
Transport-Layer Protocol	UDP	
AP	AP Bandwidth	625 kb/s
	AP Transmission Range	250 m/s
	AP Number	12

The VANET has 32 vertexes and the beginning and ending points of vehicles are randomly assigned from the 32 vertexes. The length of every line segment in the $1000\ m \times 1000\ m$ square area is $200\ m$. The length of vehicles is defined as $4\ m$ and the maximum value of number of vehicles accommodated by every line segment is 50, namely $N_{\max} = 50$. We firstly generate 10,000 movement trajectories of vehicles by simulating movement behaviors of 100 vehicles. After each vehicle is assigned the beginning and ending points, the vehicle travels along the path which has the maximum cumulative sum of speed of line segments in roads. If there are multiple candidate paths with maximum cumulative sum of speed, the vehicle is randomly assigned a path in the candidate paths. All vehicles drive into line segments with the maximum value allowed by the line segments. When the vehicles travel on line segments, they have behaviors of acceleration, deceleration and stop according to the following model [22] where $T = 1\ s$ and $\eta = 0.1$. For instance, when the speed of a vehicle is decreased/increased, the vehicle immediately moves in terms of the new speed during the assigned time period, namely the process of acceleration and deceleration is not simulated. When the number of vehicles on line segments reaches to the maximum value, the new vehicles which drive into the line segments stay at the entrance of line segments. The speed of vehicles is randomly regulated in order to simply simulate acceleration and deceleration. We make use of the 10,000 movement trajectories are used to extract movement patterns of vehicles. We also generate 300 movement trajectories of vehicles. Before the vehicles travel, they are randomly assigned the beginning and ending points and travel in terms of the movement patterns. When the vehicles arrive at the assigned ending points, they use the ending points as the new beginning points, their new ending points are re-assigned and continue to travel in terms of the movement patterns corresponding to new beginning and ending points. The 300 mobile nodes in the VANET move in terms of the 300 movement trajectories.

The media server has 30 Mb/s bandwidth and only stores a video and the length of the video is 600 s. The video is divided into 20 chunks where the length of each chunk is 30 s. The 10,200 user viewing logs are generated according to the statistics of characteristics of user playback behaviors in [23]. The 10,000 viewing logs are considered as historical playback traces and are used to extract the playback patterns of users. The 200 mobile nodes join the system following a Poisson distribution from the simulation time 0 s to 300 s and play the video content according to the 200 viewing logs. When the mobile nodes which have joined the system finish the playback for the whole video, they quit the system. The rate of video playback of mobile nodes is 128 kb/s. In PMCV, the value of threshold T_{Hm} is set to 0.5; The values of λ and θ are set to 50 and 60, respectively; The influence factor μ of time period of exchanging information between server and broker nodes is set to 1; m is the degree of fuzziness and still is set to 2; The value of α is set to 0.5. Moreover, the value of threshold V in M-PMCV is set to 0.3.

4.2 Performance Evaluation

The performance of M-PMCV is compared with PMCV according to startup delay, packet loss rate (PLR), throughput and video quality, respectively.

(1) Startup delay: The time span from the time that the nodes send the request messages to the time that the video providers receive the request messages is video lookup delay; The time span from the time that the video providers send the video data to the nodes receive the first video data is transmission delay of video data. The startup delay is composed of delay of video lookup and transmission. The average startup delay (ASD) is the mean value of startup delay of all nodes during a time span T or the number of request nodes.

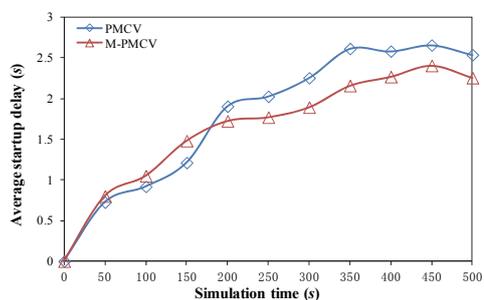


Fig. 8. ASD variation against simulation time.

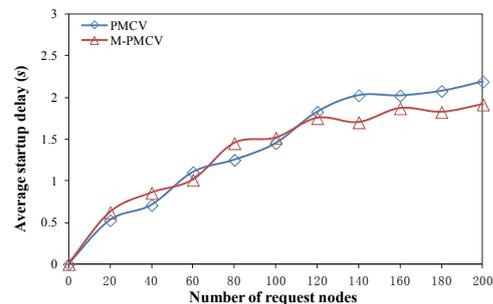


Fig. 9. ASD variation against number of request nodes

Figs. 8 and 9 show variation of the two curves corresponding to ASD results of PMCV and M-PMCV with the increase in simulation time and number of request nodes, where $T = 50$ s and the number of request nodes is 20. As Fig. 8 shows, the blue curve corresponding to M-PMCV's ASD results has a fast rise from $t = 0$ s to $t = 200$ s, experiences a relatively slow rise from $t = 250$ s to $t = 450$ s and finally has a fall from $t = 450$ s to $t = 500$ s. PMCV's curve has a relatively slow rise from $t = 0$ s to $t = 150$ s, keeps fast increase from $t = 150$ s to $t = 350$ s and starts to decrease from $t = 450$ s to $t = 500$ s

after a slight fluctuation from $t = 350$ s to $t = 450$ s. M-PMCV's red curve is higher than that of PMCV from $t = 0$ s to $t = 150$ s, but M-PMCV's red curve is lower than that of PMCV and has lower increment and peak value than PMCV from $t = 200$ s to $t = 500$ s.

In Fig. 9, the red curve corresponding to M-PMCV's ASD results has a fast rise with the increase of number of request nodes from 0 to 80 and keeps a slow increase with the increase in number of request nodes from 80 to 200. PMCV's ASD results always keep a fast rise trend with the increase in number of request nodes from 0 to 200. M-PMCV's red curve is higher than that of PMCV with the increase in number of request nodes from 0 to 100, but the red curve corresponding to M-PMCV's ASD results is lower than that of PMCV and has lower increment and peak value than PMCV with the increase in number of request nodes from 120 to 200.

The delay of video lookup and transmission determines startup delay of request nodes. PMCV constructs node communities in terms of similarity of playback and mobility between nodes. Because M-PMCV and PMCV have the same measurement method of similarity of playback behaviors (The measurement method of mobility similarity of M-PMCV and PMCV is different) and the nodes which have similar playback and dissimilar mobility become associate members in communities, M-PMCV and PMCV construct the same node communities. In other words, the number of ordinary and associate members in the constructed communities of M-PMCV and PMCV may be different, so the request nodes in the constructed communities of M-PMCV and PMCV may search to the different video providers. Therefore, the transmission delay of request messages and video data determines the difference levels of ASD performance of M-PMCV and PMCV. At the beginning of simulation, the small number of nodes start to join the system and request video content, they only obtain the video resources from the server. Therefore, the ASD results of M-PMCV and PMCV keep the low levels. With increasing number of request nodes, the node communities are consecutively constructed in M-PMCV and PMCV and the nodes which have joined into the communities fetch the video resources from video providers in communities. PMCV makes use of the AP passed by vehicles to denote the movement trajectories of vehicles, so that the deployment strategy of APs in VANET determines representation accuracy of movement behaviors of vehicles. The deployment strategy of APs relies on coverage range and deployment cost of APs. As Fig. 7 shows, the 12 APs can achieve the whole coverage for the $1000\text{ m} \times 1000\text{ m}$ square area, but it is difficult to ensure the accurate representation for the movement behaviors of vehicles. For instance, when a vehicle A moves from S_{E0} to S_{G0} , it accesses to the three APs: AP_1 , AP_4 and AP_2 . If another vehicle B moves from S_{F0} to S_{H0} , it also accesses to the three APs: AP_1 , AP_4 and AP_2 . Therefore, PMCV does not accurately describe the movement trajectories of vehicles. Moreover, PMCV makes use of the Markov process to predict the movement behaviors of vehicles in the future and estimates similarity of movement behaviors of vehicles in terms of the predicted and historical movement behaviors of vehicles. Because PMCV does not consider traffic of roads and characteristics of movement behaviors of vehicles, the accuracy levels of predicted movement behaviors based on the Markov process is difficultly ensured. The inaccurate representation of movement behaviors also results in low estimation accuracy of mobility similarity between vehicles, which brings severely negative influence for community construction and provider selection. For instance, the low estimation accuracy of mobility similarity between vehicles causes errors in community construction, so that

some nodes may be the associate members in communities rather than the ordinary members; If the relationship between the two nodes n_i and n_j are mistakenly measured to the ordinary members and n_i is selected as the video provider of n_j , the transmission performance of request message and video data is negatively influenced by the dissimilar mobility of n_i and n_j . In particular, the transmission delay of request message and video data may be lengthened with increasing movement speed of n_i and n_j . M-PMCV makes use of the refined geographical location of vehicles to estimate subsection relationship between vehicles and line segments in roads. The movement behaviors of vehicles can be accurately described by the line segments in roads in terms of the calculated subsection relationship. By modeling road traffic and analysis of historical movement trajectories to extract movement patterns of vehicles, M-PMCV can accurately recognize the movement patterns of vehicles according to characteristics of movement behaviors of vehicles and makes use of the existing movement trajectories to calculate similarity of movement behaviors between vehicles. The accurate representation of movement behaviors and the extraction and recognition of movement patterns can ensure the high accuracy of similarity estimation of movement behaviors between vehicles. The accurate estimation of similarity of movement behaviors between vehicles reduces the risk of transmission delay jitter caused by variation of transmission paths of request messages and video data with high probability. M-PMCV can efficiently decrease the transmission delay of request messages and video data. Therefore, the ASD results of M-PMCV are better than those of PMCV with increasing simulation time and number of request nodes.

Packet loss rate (PLR): The packet loss rate is the ratio between number of video data received by video requesters and total number of video data sent by video providers. The mean value of PLR of all nodes during a time interval T or with the increase in the number request nodes is defined as the average PLR.

Figs. 10 and 11 show the variation of PLR results of PMCV and M-PMCV with the increase in simulation time and number of request nodes, where $T = 50$ s and the number of request nodes is 20. As Fig. 10 shows, PMCV's blue curve keeps a fast rise trend from $t = 0$ s to $t = 350$ s and has a relatively severe fluctuation from $t = 350$ s to $t = 500$ s. M-PMCV's red curve experiences a fast rise trend from $t = 0$ s to $t = 350$ s and has a slight fluctuation from $t = 350$ s to $t = 500$ s. M-PMCV's PLR results are higher than those of PMCV at $t = 50$ s, $t = 100$ s and $t = 250$ s, but M-PMCV's PLR results are less than those of PMCV during other time period. M-PMCV's peak value (0.435 at $t = 500$ s) is less than that of PMCV (0.507 at $t = 450$ s). As Fig. 11 shows, PMCV's blue curve experiences a fast rise with increasing number of request nodes from 0 to 100 and keeps slow increase with increasing number of request nodes from 100 to 200. M-PMCV's red curve has a fast rise with increasing number of request nodes from 0 to 120 and slowly increases with increasing number of request nodes from 140 to 200 after a slight fall. When the number of request nodes is 20, 40 and 180 respectively, M-PMCV's PLR results are larger than those of PMCV; In other cases, M-PMCV's PLR results are less than those of PMCV. M-PMCV's peak value is less than that of PMCV when the number of request nodes is 200.

The 200 mobile nodes join the system and request video content from $t = 0$ s to $t = 300$ s following the Poisson distribution. Initially, the small number of nodes start to re-

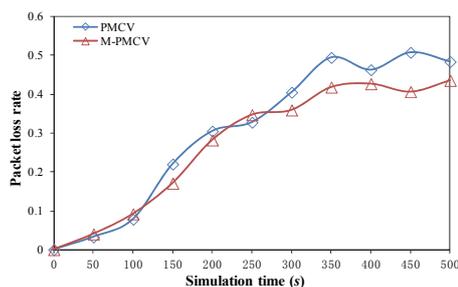


Fig. 10. PLR variation against simulation time.

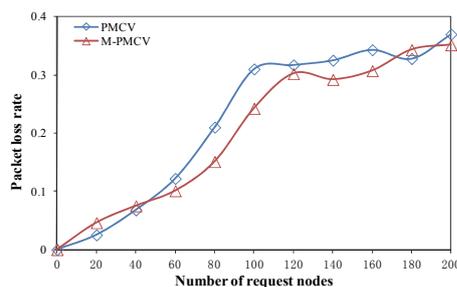


Fig. 11. PLR variation against number of request nodes.

quest video, so the VANET has relatively enough bandwidth to meet the demand of low video traffic. There are the low PLR values in PMCV and M-PMCV. With the fast increase in the number of request nodes, the increasing video traffic consumes large amount of network bandwidth in VANET, which results in the network congestion. The network congestion leads to large number of video data loss. For instance, as Fig. 10 and Fig. 11 show, the PLR results of PMCV and M-PMCV quickly increase with the increase in simulation time and number of request nodes. PMCV and M-PMCV rely on the similar movement behaviors between vehicles to achieve traffic offloading in underlying network, which reduces traffic load of core networks. PMCV makes use of the APs passed by vehicles to describe movement trajectories of vehicles. The AP-based representation of movement trajectories does not accurately describe the movement behaviors of vehicles, which brings the error between realistic and described movement trajectories. The error leads to the low estimation accuracy of similarity of movement behaviors between vehicles, which results in the increase of probability that the video requesters select the video providers with dissimilar movement behaviors. The dissimilar movement behaviors may lengthen the geographical distance between video requesters and providers, which increases the risk of video data loss. Moreover, the dissimilar movement behaviors also lead to increasing number of relay nodes between video requesters and providers, which increases consumption of network bandwidth and easily causes network congestion. The high-level network congestion leads to the large amount of video data loss. By estimation of subsection relationship between geographical location of vehicles and line segments in roads, M-PMCV makes use of the line segments in roads to accurately represent the movement trajectories of vehicles, which efficiently supports accurate estimation for similarity of movement behaviors between vehicles. By estimation of road traffic and analysis for the historical movement trajectories of vehicles to extract movement patterns of vehicles, M-PMCV can efficiently recognize movement patterns of vehicles and accurately measure similarity of movement behaviors between vehicles. The accurate measurement for similarity of movement behaviors between vehicles can make the video requesters selecting the video providers which have similar movement behaviors with high probability, which can efficiently decrease number of relay nodes in the transmission paths between video requesters and providers to support video traffic offloading in underlying network. Therefore, M-PMCV has relatively lower levels of network congestion than those of PMCV, so that the PLR results of M-PMCV are better than those of PMCV.

Average throughput: The scale of video data received by the request nodes can be defined as the product between number and size of received video data. The ratio between scale of video data received by all request nodes and a certain time period T is used to denote the total system throughput; The ratio between scale of video data received by the request nodes and the corresponding number of request nodes is used to denote the average throughput.

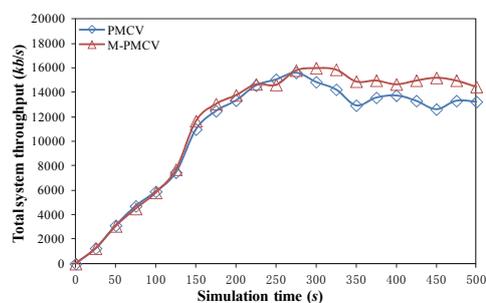


Fig. 12. Total system throughput variation against simulation time.

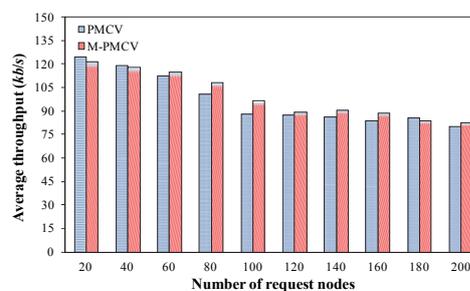


Fig. 13. Average throughput variation against number of request nodes.

Fig. 12 shows the variation of total system throughput of PMCV and M-PMCV with increasing simulation time, where $T = 25$ s. In Fig. 12, the blue curve corresponding to PMCV's results has a fast increase from $t = 0$ s to $t = 275$ s, starts to fall from $t = 275$ s to $t = 350$ s and keeps a relatively severe jitter from $t = 350$ s to $t = 500$ s. PMCV's blue curve reaches the peak value (15613.44 kb/s) at $t = 275$ s. The red curve corresponding to M-PMCV's results quickly increases from $t = 0$ s to $t = 325$ s and keeps a slight fluctuation from $t = 350$ s to $t = 500$ s after a transitory fall from $t = 325$ s to $t = 350$ s. M-PMCV's red curve reaches the peak value (15999.36 kb/s) at $t = 300$ s. M-PMCV has the longer time period of increasing throughput than that of PMCV and also has the larger peak value than that of PMCV. Fig. 13 shows the variation of average throughput of PMCV and M-PMCV with increasing number of request nodes. The blue bars corresponding to PMCV's results keep a decrease trend with the increase in number of request nodes from 20 to 100 and have a slight fluctuation with the increase in number of request nodes from 120 to 200. The red bars corresponding to M-PMCV's results also have a fall process with the increase in number of request nodes from 20 to 100 and keep a slight fluctuation with the increase in number of request nodes from 120 to 200. Although M-PMCV's average throughput values are lower than those of PMCV when the number of request nodes is 20, 40 and 180 respectively, the performance of M-PMCV's average throughput is better than that of PMCV in other cases.

In PMCV and M-PMCV, the mobile nodes join into the system and request video content following a Poisson distribution from $t = 0$ s to $t = 300$ s. With increasing number of request nodes, the video traffic in network fast increases (The two curves corresponding to PMCV and M-PMCV keep the fast rise trend from $t = 0$ s to $t = 275$ s). The increase in video traffic also brings the heavy load for the network and causes local network congestion. After $t = 300$ s, the 200 request nodes search and fetch video content, so that the network congestion caused by huge video traffic results in large-scale loss of video data. The total system throughput of PMCV and M-PMCV decreases and keeps

fluctuation from $t = 300$ s to $t = 500$ s. Fig. 13 also shows the increasing video traffic brings the negative influence for the average throughput of every node. When the number of request nodes fast increases, the average throughput of PMCV and M-PMCV keeps the fast decrease. If the transmission paths of video data between video requesters and providers have many relay nodes and multiple transmission paths go through the same relay nodes, the bandwidth of relay nodes is quickly consumed, which results in the local network congestion and causes large amount of video data loss. PMCV and M-PMCV want to make use of the similarity of movement behaviors between video requesters and providers to achieve traffic offloading in underlay network. This is as the similar movement behaviors between video requesters and providers can reduce the geographical distance and decrease the number of relay nodes in transmission paths. PMCV employs the AP-based representation method of movement trajectories of vehicles, which difficultly ensures the movement behaviors of vehicles be accurately described. Moreover, PMCV models the movement behaviors to a Markov process, which predicts the movement behaviors of vehicles in the future. The inaccurate representation of movement behaviors of vehicles increases the error of prediction results of movement behaviors between vehicles with high probability. On the other hand, PMCV neglects the characteristics of movement behaviors of vehicles, which results in the low prediction accuracy of movement behaviors. In PMCV, the video requesters select the providers which have the dissimilar movement behaviors with high probability, which easily results in the video transmission with long geographical distance. Therefore, the severe network congestion in PMCV results in large amount of video data loss and low system throughput and average throughput. M-PMCV employs the representation method of movement trajectories of vehicles based on line segments in roads. The accurate representation for movement trajectories of vehicles can efficiently support the accurate estimation of similarity of movement behaviors between vehicles. By estimation of road traffic and investigation for the characteristics of movement behaviors of vehicles, M-PMCV extracts and recognizes the movement patterns of vehicles and further calculates the similarity of movement behaviors between vehicles. Therefore, in M-PMCV, the video requesters can select the nodes which have similar movement behaviors as the providers with high probability, which relieves network congestion level and reduces number of lost video data. Therefore, the throughput performance of M-PMCV is better than that of PMCV.

Video quality: The Peak Signal-to-Noise Ratio (PSNR) is used to denote the video quality of PMCV and M-PMCV [24]. The PSNR can be defined as:

$$PSNR = 20 \cdot \log_{10} \left(\frac{MAX_Bit}{\sqrt{(EXP_Thr - CRT_Thr)^2}} \right) \quad (14)$$

where EXP_Thr and CRT_Thr are the expected average throughput and actual measured throughput, respectively; MAX_Bit denotes the average bitrate of the video stream. The values of MAX_Bit and EXP_Thr are set to 128 kb/s in terms of the simulation settings. The PSNR of every request nodes is considered as average PSNR.

Fig. 14 shows the PSNR variation of PMCV and M-PMCV with increasing number of request nodes. The blue bars corresponding to the average PSNR values of PMCV

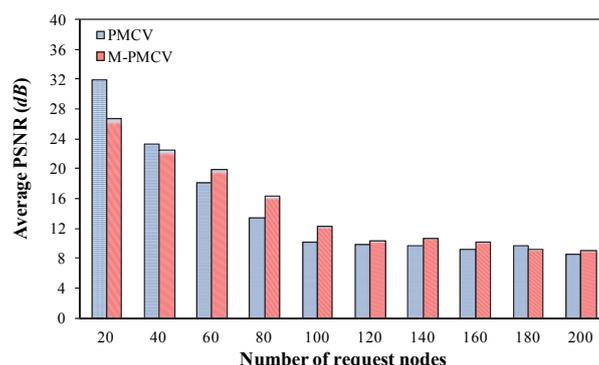


Fig. 14. Average PSNR variation against number of request nodes.

have a fast fall with the increase in number of request nodes from 0 to 100 and keep a slight jitter from 120 to 200. The red bars of M-PMCV also have the same variation with PMCV, namely slight jitter from 120 to 200 after a fast decrease from 0 to 100. Although M-PMCV's average PSNR values are lower than those of PMCV when the number of request nodes is 20, 40 and 180 respectively, M-PMCV's average PSNR values are higher than those of PMCV in other cases.

The average PSNR indicates that the experience quality of each request node in the process of watching video. The larger the PSNR value is, the better the watched quality of users is. The performance of PSNR is related to playback rate and PLR. The high playback rate denotes the request nodes need to receive large amount of video data to support smooth and high-definition playback. However, the high playback rate requires more network bandwidth, which increases the network congestion levels. The high PLR indicates that the request nodes do not receive enough video data to support smooth playback, which brings severely negative influence for the viewing quality of users. In PMCV, the low accuracy for the similarity estimation of movement behaviors between vehicles results in relatively high PLR, so that PMCV's PSNR values quickly decrease and finally keep fluctuation at low levels. In M-PMCV, the accurate similarity estimation of movement behaviors between vehicles efficiently supports video delivery with close geographical distance to reduce PLR. Although M-PMCV's PSNR results also quickly decrease, the decrement is relatively low and has a fluctuation at relative high levels. Therefore, the performance of M-PMCV's average PSNR is better than that of PMCV.

5. CONCLUSIONS

In this paper, we propose a novel mobility similarity measurement method to increase the performance of community-based video delivery in VANETs (MSMM). MSMM refines vehicle geographical location and uses the line segments in roads to describe the movement trajectories of vehicles by estimation of subjection relationship between vehicle location and roads. MSMM constructs an estimation model of traffic of road in terms of the hydromechanics and the vehicle following model and extracts the movement patterns of vehicles by analysis for the historical movement trajectories of vehicles to calculate traffic of roads. MSMM recognizes the movement patterns of vehi-

cles in terms of current movement trajectories of vehicles and uses the recognized patterns to calculate the similarity of movement behaviors between vehicles. MSMM is used to replace estimation component of similarity of node mobility in our previous work PMCV and construct a new video sharing solution (called “M-PMCV”) in VANET. The simulation results show that M-PMCV achieves lower startup delay, lower packet loss rate, high throughput and high video quality than PMCV.

ACKNOWLEDGMENT

This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant Nos. 61501216, 61402303, 61572246, U1604155 and U1404611, the Project of Beijing Municipal Commission of Education under Grant KM201510028016, the Open Foundation of State key Laboratory of Networking and Switching Technology (Beijing University of Posts and Telecommunications) under Grant no. SKLNST-2016-2-12, the Science and Technology Key Project of Henan province under Grant No. 182102210105, the Program for Science and Technology Innovation Talents in the University of Henan Province under Grant No. 16HASTIT035, the Program for Science and Technology Innovation Outstanding Talents in the Henan Province under Grant no. 184200510011, the Inter-Governmental Science and Technology Cooperation Project under Grant No. 2016YFE0104600 and the IRTSTHN Grant No. 18IRTSTHN014.

REFERENCES

1. S. Chen and J. Zhao, “The requirements, challenges and technologies for 5G of terrestrial mobile telecommunication,” *IEEE Communications Magazine*, Vol. 52, 2014, pp. 36-43.
2. Y. Ye, S. Ci, N. Lin, and Y. Qian, “Cross-layer design for delay-and energy-constrained multimedia delivery in mobile terminals,” *IEEE Wireless Communications*, Vol. 21, 2014, pp. 62-69.
3. L. Sun, H. Shan, A. Huang, L. Cai, and H. He, “Channel allocation for adaptive video streaming in vehicular networks,” *IEEE Transactions on Vehicular Technology*, Vol. 66, 2016, pp. 734-747.
4. N. Kumar, J. J. P. C. Rodrigues, and S. Zeadally, “On efficient and scalable support of continuous queries in mobile peer-to-peer environments,” *IEEE Transactions on Industrial Electronics*, Vol. 62, 2015, pp. 7892-7900.
5. C. Xu, F. Zhao, J. Guan, H. Zhang, and G.-M. Muntean, “QoE-driven user-centric VoD services in urban multihomed P2P-based vehicular networks,” *IEEE Transactions on Vehicular Technology*, Vol. 62, 2013, pp. 2273-2289.
6. C. Xu, S. Jia, L. Zhong, H. Zhang, and G.-M. Muntean, “Ant-inspired mini-community-based solution for video-on-demand services in wireless mobile networks,” *IEEE Transactions on Broadcasting*, Vol. 60, 2014, pp. 322-335.
7. H. Shen, Z. Li, Y. Lin, and J. Li, “SocialTube: P2P-assisted video sharing in online social networks,” *IEEE Transactions on Parallel and Distributed Systems*, Vol. 25, 2014, pp. 2428-2440.
8. C. Xu, S. Jia, L. Zhong, and G.-M. Muntean, “Socially aware mobile peer-to-peer

- communications for community multimedia streaming services,” *IEEE Communications Magazine*, Vol. 53, 2015, pp. 150-156.
9. C. Xu, S. Jia, M. Wang, L. Zhong, H. Zhang, and G.-M. Muntean, “Performance-aware mobile community-based VoD streaming over vehicular ad hoc networks,” *IEEE Transactions on Vehicular Technology*, Vol. 64, 2015, pp. 1201-1217.
 10. O. T. Cruces, M. Fiore, and J. M. B. Ordinas, “Cooperative download in vehicular environments,” *IEEE Transactions on Mobile Computing*, Vol. 11, 2012, pp. 663-678.
 11. J. Yoon, B. Noble, M. Liu, and M. Kim, “Building realistic mobility models from coarse-grained traces,” in *Proceedings of ACM International Conference on Mobile Systems, Applications and Services*, 2006, pp. 177-190.
 12. M. Kim, D. Kotz, and S. Kim, “Extracting a mobility model from real user traces,” in *Proceedings of IEEE INFOCOM*, 2006, pp. 1-13.
 13. S. Jia, C. Xu, J. Guan, H. Zhang, and G.-M. Muntean, “A novel cooperative content fetching-based strategy to increase the quality of video delivery to mobile users in wireless networks,” *IEEE Transactions on Broadcasting*, Vol. 60, 2014, pp. 370-384.
 14. J.-G. Lee, J. Han, and K.-Y. Whang, “Trajectory clustering: a partition-and-group framework,” in *Proceedings of ACM SIGMOD*, 2007, pp. 593-604.
 15. J.-K. Lee and J. C. Hou, “Modeling steady-state and transient behaviors of user mobility: formulation, analysis, and application,” in *Proceedings of ACM International Symposium on Mobile Ad Hoc Networking and Computing*, 2006, pp. 85-96.
 16. J. Lim, H. Yu, K. Kim, M. Kim, and S.-B. Lee, “Preserving location privacy of connected vehicles with highly accurate location updates,” *IEEE Communications Letters*, Vol. 21, 2017, pp. 540-543.
 17. Y. Zhang, C. C. Tan, F. Xu, H. Han, and Q. Li, “VProof: Lightweight privacy-preserving vehicle location proofs,” *IEEE Transactions on Vehicular Technology*, Vol. 64, 2015, pp. 378-385.
 18. O. Aloquili, A. Elbanna, and A. Al-Azizi, “Automatic vehicle location tracking system based on GIS environment,” *IET Software*, Vol. 3, 2009, pp. 255-263.
 19. H. Staras and S. N. Honickman, “The accuracy of vehicle location by trilateration in a dense urban environment,” *IEEE Transactions on Vehicular Technology*, Vol. 21, 1972, pp. 38-43.
 20. J. Du and M. J. Barth, “Next-generation automated vehicle location systems: Positioning at the lane level,” *IEEE Transactions on Intelligent Transportation Systems*, Vol. 9, 2008, pp. 48-57.
 21. C.-H. Ku and W.-H. Tsai, “Obstacle avoidance in person following for vision-based autonomous land vehicle guidance using vehicle location estimation and quadratic pattern classifier,” *IEEE Transactions on Industrial Electronics*, Vol. 48, 2001, pp. 205-215.
 22. D. Helbing and B. Tilch, “Generalized force model of traffic dynamics,” *Physical Review E*, Vol. 58, 1998, pp. 133.
 23. A. Brampton, A. MacQuire, I. Rai, N. J. P. Race, L. Mathy, and M. Fry, “Characterising user interactivity for sports video-on-demand,” in *Proceedings of ACM NOS-SDAV*, 2007, pp. 1-6.
 24. S.-B. Lee, G.-M. Muntean, and A. F. Smeaton, “Performance-aware replication of distributed pre-recorded IPTV content,” *IEEE Transactions on Broadcasting*, Vol. 55, 2009, pp. 516-526.



Shi-Jie Jia (贾世杰) received Ph.D. degree in Communications and Information System from BUPT in March 2014. He is an Associate Professor at the Academy of Information Technology, Luoyang Normal University, Henan, China. His research interests include next generation Internet technology, wireless communications, and peer-to-peer networks.



Rui-Ling Zhang (张瑞玲) received the M.S. degree from Northwestern Polytechnical University, Xi'an, China, in 2007. She is currently a Professor at Luoyang Normal University. Her research interests include intelligent information processing, data mining and rough set.



Tian-Yin Wand (王天银) received the Ph.D. degree in Cryptography from Beijing University of Post and Telecommunications in 2010. He is currently a Professor at Luoyang Normal University. His research interests include cryptography and information security.



Lu-Jie Zhong (袁璐洁) is the corresponding author, email: zl-jict@gmail.com. She received her Ph.D. degree from Institute of Computing Technology, Chinese Academy of Sciences, in July 2013. She is an Associate Professor in the Information Engineering College, Capital Normal University, China. Her research interests are in the areas of communication networks, computer systems and architecture, mobile Internet technology, the Internet of things, and vehicular networks.



Ming-Chuan Zhang (张明川) received his Ph.D. degree from Beijing University of Posts and Telecommunications in July 2014. He is an Associate Professor in Henan University of Science and Technology. His research interests include bio-inspired networks, Internet of Things, future Internet and computer security.