
A Scheme of Selecting Vehicles to Assist Download Based on WebGIS for VANET

Qibin Zhou¹, Qinggang Su^{2,†} and Peng Xiong^{3,*}

¹*College of International Education, Shanghai Jian Qiao University, Shanghai China*

²*Kaiserslautern Kolleg für Intelligente Produktion, Shanghai Dianji University, Shanghai China*

³*School of Electronics and Information, Shanghai Dianji University, Shanghai China*

E-mail: zqbin@gench.edu.cn; suqg@sdju.edu.cn; xiongp@sdju.edu.cn

**Corresponding Author*

†Co-first Author

Received 26 July 2021; Accepted 28 August 2021;
Publication 03 January 2022

Abstract

The assisted download is an effective method solving the problem that the coverage range is insufficient when Wi-Fi access is used in VANET. For the low utilization of time-space resource within blind area and unbalanced download services in VANET, this paper proposes an approximate global optimum scheme to select vehicle based on WebGIS for assistance download. For WebGIS, this scheme uses a two-dimensional matrix to respectively define the time-space resource and the vehicle selecting behavior, and uses Markov Decision Process to solve the problem of time-space resource allocation within blind area, and utilizes the communication features of VANET to simplify the behavior space of vehicle selection so as to reduce the computing complexity. At the same time, Euclidean Distance(Metric) and Manhattan Distance are used as the basis of vehicle selection by the proposed scheme so that, in the case of possessing the balanced assisted download

Journal of Web Engineering, Vol. 21_2, 337–364.

doi: 10.13052/jwe1540-9589.2129

© 2022 River Publishers

services, the target vehicles can increase effectively the total amount of user downloads. Experimental results show that because of the wider access range and platform independence of WebGIS, when user is in the case of relatively balanced download services, the total amount of downloads is increased by more than 20%. Moreover, WebGIS usually only needs to use Web browser (sometimes add some plug-ins) on the client side, so the system cost is greatly reduced.

Keywords: Assisted download, WebGIS, delay-tolerant networking, VANET.

Introduction

Recent years, the vehicle with the help of the technology of IEEE 802.11 and DSRC to access Internet through roadside infrastructure has become an important means for in-vehicle users to achieve information interaction with Internet as Wi-Fi access is deployed on a large scale. Compared to the method utilizing 3G/4G to access Internet, Wi-Fi access can give the lower cost and better service to the high-speed mobile ends. Now, applying Wi-Fi access to VANET has been widely approved by the researchers in the whole world. But the insufficient covering range of Wi-Fi access prevents from its universal in VANET.

By the vehicles which run in the same direction or opposite direction carrying data to increase the download amount of target vehicle, this method of assisted download can solve effectively the problem that the insufficient covering range of Wi-Fi access prevents from its extending in VANET [1]. But it cannot solve the problem that the utilizing rate of time-space resources is low and the download services are unbalanced.

In VANET, in the case of multi-user requests, the scheme helping the target vehicle running in the blind area(BA) to select the assisted download vehicles is one of the key factors to decide to the system throughput. The goal of vehicle selection scheme is to maximize the total amount of user download on the premise that all the target vehicles can evenly possess the assisted download services. This scheme is facing with two problem to solve: (1) Since the assisting vehicles don't enter into the communication area (CA) of an access point (AP) at the same time, during these vehicle staying in CA, AP must decide whether to used them as the assisting vehicles to carry data and the number of carried data, and to provide the assisted download services for which target vehicles; (2) The actual transmission between the

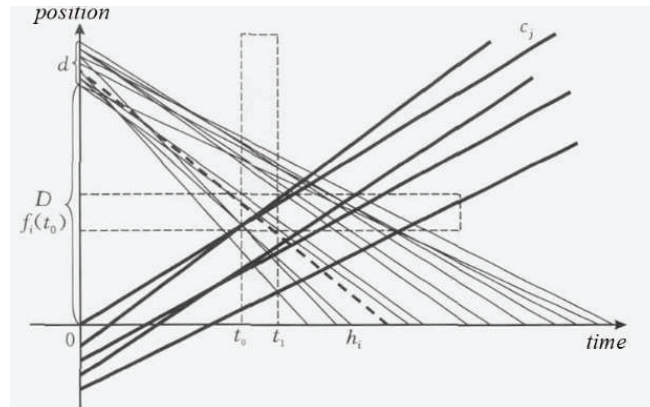


Figure 1 Schematic diagram of the encounter between the assistance vehicle and the target vehicle.

assisting vehicle and the target vehicle occurs a period of time after AP makes a decision. So, all decisions are based on predictions, which may lead to uncertainty in the results.

The basis of selecting the vehicles is based on the existing information (the speed and time enter into AP etc.) from the registered vehicle to predict the time and location where the target vehicle and the assisting vehicle makes communication within BA to select the assisting vehicles with the non-overlapping communication area. As shown in Figure 1, the horizontal axis denotes time, the vertical axis denotes position, the thin line denotes the assisting vehicle, and the thick line denotes the target vehicle. The target vehicle sets off from the present AP ($position = 0$). The assisting vehicle sets off from the next AP ($position = D$), and it meets with the different targets in the different time and location. t_0 and t_i denote respectively the starting time and ending time when the target vehicle c_j and assisting vehicle h_i communicates.

The vehicles don't enter into CA of AP at the same time, so if every decision from AP is only consider the vehicles which are in CA at that time, the result of selecting vehicle is local optimum. However, due to the superimposition of wireless communication conflict field in the case of a single channel, the made decision at the time t can influence the made decision at the time $t + 1$. Obviously, this local optimum may not necessarily equal the global optimum.

In addition to the maximum throughput of the system, maintaining the fairness of data obtained by the target vehicle, that is, each target

vehicle should possess the assisted download service as equal as possible, is another factor influencing the decision for selecting vehicle. In the existing researches, the scheme of selection vehicle with cycle in sequence [2] can make each target vehicle obtain the download services in a balanced method. But the utilizing rate of BA is low so as to reduce the amount of user download.

For the above problem, this paper proposes the scheme to select vehicle for assisted download with the approximate global optimum in VANET. The proposed scheme utilizes the communication features of VANET to simply the space of selecting vehicle so as to reduce the computing complexity. Euclidean Distance (ED) and Manhattan Distance (MD) are used to determine the selection probability of different behavior so as to balance the number that each target vehicle possesses the assisted download services. The proposed scheme is proposed to improve the utilization of time-space resource within BA on the premise that CA is not overlapped in each procedure of the assisted download so as to improve the amount of user download and meet the download requirement for mass data.

1 Relate Work

Depending on application environment, vehicle selection scheme for assisted download can be divide into two types: urban road environment and highway environment.

For the urban road environment, due to complex road conditions and large changes in vehicle speed, some vehicle selection schemes to assist download are researched based on the prediction for vehicle travel route, such as the references [3–5]. The reference [3] uses a multi-level Markov chain to predict the next location of vehicle based on the current travel route of vehicle so as to select the assisting vehicle to provide the assisted download services for the designated target vehicle. The references [4, 5] propose to use the road topology to predict vehicle travel route so as to make vehicle selection scheme. The reference [6] proposes the vehicle selection scheme which predicts the possible travel route of vehicle according to its history routes, selecting the vehicle with high encounter probability as the assisting vehicle. In addition to using vehicle travel route as the basis for the selection scheme, the references [7, 8] also propose to use the deployment of AP nodes as the basis for vehicle selection. The goal of AP deployment scheme in Alpha Coverage [7] is to use as few AP as possible to provide users with access services. And although the scheme α -coverage reduces the number of access

points, it does not consider the quality of services provided by AP to users. For this problem, the author optimized the deployment scheme of AP on the basis of α -coverage, and innovatively proposed a measurement standard Contact Opportunity [8]. Compared with α -coverage, Contact Opportunity considers the dynamic and static information such as the coverage area of AP, the vehicle speed, the data download rate of AP, etc., so that AP can be deployed more reasonably.

For the highway environment, the schemes for vehicle selection in the urban road environment cannot be applied well in the highway environment. The reasons include: (1) It is difficult to predict the vehicle travel route in the urban road environment, so the vehicle selection schemes to assist download mainly focus on the researches on the travel route prediction and encounter probability; (2) In the highway environment, the road topology is simple. The number of intersections is much lower than urban environment, so it is possible to deploy AP at each intersection so that any two vehicles between two AP in only possible to travel in the opposite or same direction, and the two vehicles traveling in the opposite directions inevitably encounter each other. And due to the lower change rate of speed, the two vehicles traveling in the same direction are easy to predict whether they will encounter. In addition, in the highway environment, the number of vehicles per unit area is small, the speed of vehicle is fast, the change rate of speed is low, and the number of vehicles available for AP in a unit time is small. These features are not available in the urban road environment. So, the vehicle selection scheme to assist download in the highway environment focuses more on optimizing the time-space resources within BA and the vehicles. In the highway environment, the local optimum scheme is used in the transmission method of MobTorrent [4], that is, each time AP selects the vehicle which can provide the largest download in the current time based on the vehicles in CA as the assisting vehicle. The reference [6] uses the vehicle selection scheme with polling in sequence, that is, the target vehicle that makes the request first can get the services first. And the reference [9] uses the completion rate as the vehicle selection basis, that is, each time the assisting vehicle selected is the vehicle which can provide the users with the services having the highest download completion rate.

The vehicle selection schemes adopted in the above researches are almost the local optimum or the in-sequential cycle. These schemes lead to the low utilization of time-space resources within BA and the problem of unbalanced download services in the case of multiple user requests. How to select the assisting vehicle to provide the target vehicle within BA with the download

service so as to effectively improve the utilization rate of BA needs an in-depth research.

Optimization theory is one of the popular methods to solve the problem of resource allocation and task scheduling. The basic theoretical model of dynamic optimization proposed in the reference [10] is Markov Decision Process (MDP). MDP can be used to describe such a discrete-time decision-making process: the transition of system state at time $t + 1$ depends only on the system state at time t and the decision-maker's behavior, and is unrelated to the system state and the decision-maker's behavior in the period $[0, t - 1]$. This is consistent with the space-time resource allocation scheme within BA analyzed above. So, this paper proposes to use the dynamic optimization model of MDP as the basis for vehicle selection to solve the problem of the time-space resources allocation within BA.

2 Optimization Model Based on MDP

2.1 Fundamental Element

Here, a dynamic optimization model of MDP is built based on the features of time-space resources within BA in VANET. The fundamental elements required are as follows:

- (1) Two-dimensional state set $S = s(\tau, p)$ denotes the time-space state within BA, that is, the resource occupation at the location p and time τ . Here, $[0, 1]$ are used to denote free and occupied respectively. The research in this paper is based on the single-channel situation, that is, the overlap of CA during the assisted download process will definitely cause transmission conflicts. Three-dimensional state set $S = s(t, p, c)$ (the resource occupancy of channel c at location p and time t) in the more complex multi-channel case will be a future research.
- (2) The set $V(t)$ of registered vehicles at the current time.
- (3) The behavior set $A(t)$ denotes the possible vehicle selection behavior of the decision maker AP at time t , and the behavior space of AP depends on the registered vehicles in AP's communication area at the current moment.
- (4) The state transition relationship SA denotes the transition process of the system state under the influence of decision-making behavior, namely $S \times A \rightarrow S$, denoted as $s' = SA(s, a)$. s is the occupancy state of time-space resources within BA at the current moment, a is the vehicle selection behavior, and s' is the next state under the behavior a .

- (5) External variables. The highway traffic flow conforms to Poisson Distribution [11].
- (6) The revenue function $R(s, a, s')$ denotes the generated benefit from the system running under the impact of the behavior of the decision maker. In order to facilitate the elaboration of the model, this paper uses the total amount of assisted download of the target vehicle as the system revenue.

2.2 Modeling and Analysis

There are two premises for vehicle selection decision need to require: (1) each vehicle selection decision should avoid conflicts in all processes of assisted download within BA; (2) all target vehicles in the system can possess fairly the assisted download services, that is, the number of allocated services for each target vehicle is balanced. So, the decision-making process is just the allocation of time-space resources within BA, that is, to determine the time and location when the target vehicle and the selected assisting vehicle meet, so as to avoid the overlap of CA. The modeling process is divided into the following 5 steps:

- (1) Determine the state space

The premise on selecting a vehicle is that all assisted download processes don't conflict within BA. The proposed scheme uses a two-dimensional matrix to denote the time-space resources within BA. As shown in the formula (1), the horizontal direction denotes the spatial unit of BA, and the vertical direction denotes the time unit of BA. The time-space occupation of BA at time t is defined as the system state at time t . The "0" in the two-dimensional matrix denotes that the time-space unit is in the idle state, and the "1" indicates that the time-space unit is in the occupied state. The state space $S = \{s(1), s(2), \dots, s(t)\}$ of the system is defined as a time-dependent set of time-space states within BA.

$$s_t = \begin{matrix} & & p_1 p_2 \dots \dots \dots \\ & t + 1 & \left[\begin{matrix} 00001000001 \dots \\ 01010100001 \dots \\ 00101011000 \dots \\ 00011001001 \dots \\ \dots \dots \dots \end{matrix} \right] \end{matrix} \quad (1)$$

(2) Determine the behavior space

All possible vehicle selection behaviors from the decision-maker AP at time t are defined as their behavior space at time t . In order to better reflect the changes in the state of time-space resources within BA caused by the decision-maker's vehicle selection behavior, the paper sets the behavior space as a two-dimensional matrix with the same definition as the state space, as shown in the formula (2).

$$a_t^i = \begin{matrix} & & p_1 p_2 \cdots \cdots \cdots \\ t + 1 & \left[\begin{array}{l} 00000010000 \cdots \\ 00000001000 \cdots \\ 00000000100 \cdots \\ 00000000010 \cdots \\ \cdots \cdots \cdots \end{array} \right. \\ t + 2 & \\ \cdot & \\ \cdot & \\ \cdot & \end{matrix} \quad (2)$$

Assuming that the vehicle selection behavior at time t is $A(t) = \{a_t^0, a_t^1, \dots, a_t^n\}$, the service selecting an assisting vehicle to provide download for the target vehicle will cause some time-space units within BA to be occupied. So, the "1" in the two-dimensional matrix denotes that the current vehicle selection behavior will make communication in this time-space unit, and the "0" denotes that the assisting vehicle and the target vehicle will not communicate in this time-space unit. The vehicle selection behavior is the one-to-many, that is, an assisting vehicle can carry data for the different target vehicles, and transmit it to the corresponding target vehicles at different times and locations within BA. So, the location of the "1" is not necessarily continuous.

(3) State transition relationship

The decision-maker's vehicle selection behavior at time t will lead to the transition of the system state, as shown in the formula (3). Assuming that the current time-space state within BA is s_t , and the current behavior space is $A(t) = \{a_t^0, a_t^1, \dots, a_t^i\}$, and the behavior a_t^i will transfer the system state from the current s_t to the next state s_{t+1} . The next state is the sum of matrices of the current state and the vehicle selection behavior. In the solving process, if the corresponding location in the matrix s_t corresponding to the location of "1" in the matrix a_t is also "1", it denotes that the time-space unit has been occupied. If the selecting behavior a_t from the decision-maker is conflictive in this time-space unit, for the formula (3), the optimal solution in the current situation should be the behavior that makes the number of "1" in

s_{t+1} the most.

$$\begin{aligned}
 & \begin{matrix} s_t \\ \left[\begin{array}{l} 00001000001 \dots \\ 01010100001 \dots \\ 00101011000 \dots \\ 00011001001 \dots \\ \dots \end{array} \right] \end{matrix} + \begin{matrix} a_t^1 \\ \left[\begin{array}{l} 00010000000 \dots \\ 00001000000 \dots \\ 00000100000 \dots \\ 00000010000 \dots \\ \dots \end{array} \right] \\ \dots \\ a_t^i \\ \left[\begin{array}{l} 00000010000 \dots \\ 00000001000 \dots \\ 00000000100 \dots \\ 00000000010 \dots \\ \dots \end{array} \right] \end{matrix} = \begin{matrix} s_{t+1}^1 \\ \left[\begin{array}{l} 00011000001 \dots \\ 01011100000 \dots \\ 00101111000 \dots \\ 00011011001 \dots \\ \dots \end{array} \right] \\ \dots \\ s_{t+1}^i \\ \left[\begin{array}{l} 00001010001 \dots \\ 11010101000 \dots \\ 01101011100 \dots \\ 00111001011 \dots \\ \dots \end{array} \right] \end{matrix} \quad (3)
 \end{aligned}$$

(4) System operation target

In the above state transition process, the decision-maker hopes that the selected behavior can obtain the maximum benefit, and the system goal determines its benefit function. The system goal can be the maximum throughput of the system or the download completion rate of the users. The maximum throughput of the system is the ultimate goal pursued by the operators, since the more data the users obtain, the higher their benefit. From the user’s point of view, every user wants to possess the download services fairly. Moreover, the different users have different requirements for the download. So, simply using the time when the assisted download occurs as a measure of system operation goals does not fully reflect the fairness of the system.

In order to solve the above problems, the proposed scheme introduces the weight variable φ as the measuring standard. The user sets φ according to the demand, and the decision-maker provides the different services according to φ . Assuming that the benefit from the state transition caused by the decision-maker’s behavior is $R(s, a, s')$, it denotes that the adopted behavior a lead to the state transition from s to s' so as to get the maximum value of R .

$$R = \sum_{i=0}^n \varphi_j \cdot \mu \cdot (t_{i1} - t_{i2}) \times \omega \quad (4)$$

(5) Bellman Equation

The behavior space refers to all the vehicle selection behaviors in which AP selects the assisting vehicle to carry data for one or several target vehicles within its communication area. In the assisted download schemes for VANET, the behavior space corresponding to each state transition is the same, that is, no matter how the decision maker selects the assisting vehicle to carry data for the target vehicle at the previous moment, the candidate vehicles at the next moment are the same, and the behavior space of vehicle selection of AP is also the same. So, although the decision at the next moment depends on the decision result at the previous moment, the behavior space at the next moment is still independent of the decision at the previous moment.

Assuming that the available behavior space corresponding to the time-space resource state s_t within the current blind area is A_t , the possible behaviors $\{a_t^0, a_t^1, \dots, a_t^n\}$ at the time t respectively transfer the system state to $\{s_{t+1}^0, s_{t+1}^1, \dots, s_{t+1}^n\}$, and the corresponding benefit is $R_t(s_t, [\{a_t^0, a_t^1, \dots, a_t^n\}, [s_{t+1}^0, s_{t+1}^1, \dots, s_{t+1}^n]])$, a key goal of the system is as follow.

$$J = \max \left(\sum_{i=0}^n R_i(s_i, a, s_{i+1}) \right) \quad (5)$$

Obviously, this is a global optimal solution, but in fact the global optimal solution is not equal to the sum of the local optimal solutions, as in the formula (6).

$$\max \left(\sum_{i=0}^n R_i(s_i, a, s_{i+1}) \neq \sum_{i=0}^n \max R_i(s_i, a, s_{i+1}) \right) \quad (6)$$

In other words, the optimal solution from the formula (3) at time t is not necessarily the solution of the optimal solution sequence at time t , but it is possible that among all solutions from the formula (3), the solution $a_x (a_x \in A(t) \text{ and } a_x \neq a_t)$ is the solution of the optimal solution sequence at time t . So, the decision made at time t must consider its impact on the time $t + 1$. Let the sequence λ be the strategy sequence solution, and $J^\lambda(s)$ is the expected value of the objective function obtained on the state $s \in S$ on the premise of adopting λ . As shown in the formula (7).

$$J^\lambda(s_n) = R(s_n, \lambda(s_n)) \sum_{s_{n+1} \in S} P(s_{n+1}|s_n, \lambda(s_n)) J^\lambda(s_{n+1}) \quad (7)$$

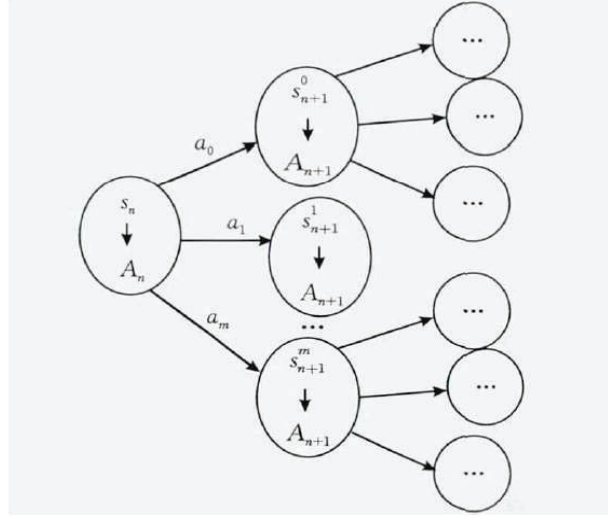


Figure 2 Bellman recursive process.

Then the solution of each moment is the formula (8).

$$\pi(s_t) = \max_{a_t^i \in A_t} \left\{ R(s_t, a_t^i, s_{t+1}) + \sum_{s_{t+1} \in S} P(s_t, a_t^i, s_{t+1}) J^\pi(s_{t+1}) \right\} \quad (8)$$

The recursive process is shown in Figure 2.

2.3 Solving Algorithm

(1) Vehicle selection behavior space

Solving the above recursive process may lead to “state space explosion”, that is, the state space increases exponentially with time. If $C = \{c_1, c_2, \dots, c_n\}$ is used to denote the set of the target vehicles, and $H = \{h_1, h_2, \dots, h_m\}$ is used to denote the set of the assisting vehicles, then the number of the target vehicles is n and the number of assisting vehicles is m . When the first assisting vehicle h_1 enters CA of AP at time t_1 , the possible decision of AP is that h_1 carries information for any combination of vehicles(e.g. c_1, c_2 , and c_5 in the set C of the target vehicles). So, the decision behavior space of AP at time t_1 is as shown in the formula (9).

$$A(h) = \{a_1, a_2, \dots, a_{2^n}\} \quad (9)$$

There are 2^n possible behaviors in $A(h)$. When other assisting vehicles enter CA of AP, AP has the same behavioral space. So, the combination of all behaviors is as follow.

$$O = (2^n)^m = 2^{n \cdot m} \quad (10)$$

Obviously, the behavior space of the formula (10) increases exponentially with the number of the target vehicles requiring to download and the number of the assisting vehicles participating in the assisting forwarding. This space is very huge. But in fact, for VANET on highway, the solution of many behavior combinations is meaningless. First, in order to prevent signal interference, when an assisting vehicle communicates with a target vehicle, the other target vehicles within the communication range of this assisting vehicle cannot receive data, and the other vehicles within the communication range of this target vehicle cannot also send data. So, the conflicting combinations in the behavior space have no practical meaning and can be excluded in the solving process. Moreover, the vehicle selection principle is to achieve the system maximum throughput. So, the communications that will not conflict with other vehicles should be a must.

(2) Simplifying behavior space

Obviously, in the case that the number n of the target vehicles is unchanged, the number m of the assisting vehicles determines the size of the behavior space and increases exponentially. The larger m , the closer the corresponding solution to the optimal solution, and the higher the time-space utilization rate for BA. But it also requires more time to calculate. So, in order to reduce the computing complexity as much as possible, the proposed scheme simplifies the behavior space according to the features of VANET.

According to the above analysis, the solution space caused by the following two special cases has no practical meaning.

(1) When CA of any two or more running vehicles in the target vehicle set C is overlapped with CA of one of the assisting vehicles, as shown in Figure 3.

In this case, AP can only select one of the target vehicles when making decisions for h_1 , that is, in the combination of formula (9), c_1 and c_2 cannot appear at the same time. Just for the case in Figure 3, the number of data sets of $A(t)$ is simplified to $(2^n - 2^{n-2})$. If other target vehicles overlap, the behavior space of $A(t)$ will be further reduced. Figure 4 is a fragment of the optimization algorithm when the collision domains between an assisting vehicle and multiple target vehicles overlap.

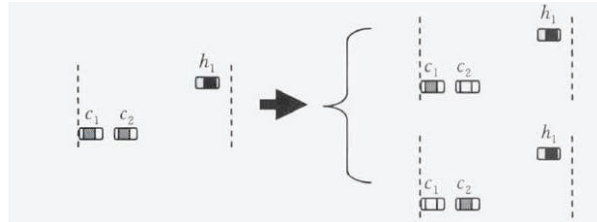


Figure 3 The communication domain between h_1 and c_1 and the communication domain between h_1 and c_2 are overlapped.

```

1.  $h$ ;
2.  $A(h)$ ;
3. for  $i:(0 \sim \max)$  in  $C$ 
4.   for  $j:(i+1 \sim \max)$  in  $C$ 
5.     if  $h_i$  collide with  $h_j$ 
6.        $A(h) = \text{remove}(A(h), i, j)$ ;
7.     end
8.   end
9. end
10. return  $A(h)$ ;

```

Figure 4 Multi-target vehicle optimization algorithm.

For any assisting vehicle h , first initializing the behavior space of h (line 2), and putting all possible combinations into the set $A(h)$. In the initial case, the number of elements in the set $A(h)$ is 2^n . For any given assisting vehicle h , using two cycles (lines 3 and 4) to traverse the target vehicle set C . According to the scheme proposed in the reference [12], it is possible to calculate whether the collision domains between any two target vehicles and the assisting vehicle overlap (line 5). If the cases are met, deleting the combination containing two assisting vehicles in the set (line 6). Repeating this operation on lines 8–9 until all combinations are checked. Line 10 returns the optimized set $A(h)$.

(2) When CA of any two or more running assisting vehicles in the assisting vehicle set H is overlapped with CA of one of the target vehicles, as shown Figure 5.

Before making a decision on each vehicle, AP must decide for which target vehicles the current assisting vehicle can provide services based on the previously selected vehicle. So, each target vehicle establishes a list of assisting vehicles to record which assisting vehicles provide services for

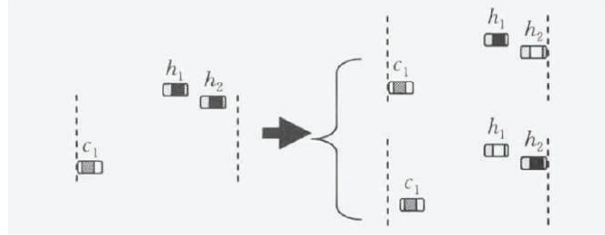


Figure 5 The communication area between h_1 and c_1 and the communication domain between h_2 and c_1 are overlapped.

```

1.  $h$ ;
2.  $A(h)$ ;
3. for each  $i$  in  $C$ 
4.   for each  $j$  in  $List(i)$ 
5.     if  $h$  collide with  $j$ 
6.        $A(h) = \text{remove}(A(h), i)$ ;
7.     end
8.   end
9. end
10. return  $A(h)$ ;

```

Figure 6 Multi-assisting vehicle optimization algorithm.

itself, and use this list as the basis for the current decision. As shown in Figure 5, since the communication domain of h_1 and c_1 overlap with the communication domain of h_2 and c_1 , AP must first check whether there is h_1 in the service list of c_1 when deciding whether h_2 provides services for c_1 . If so, in the data set $A(t)$, all options containing h_2 will be deleted. By this way, the behavior space of $A(t)$ is reduced.

Figure 6 is a fragment of the optimization algorithm that simplifies the behavior space in the case that overlaps the collision domain between multiple assisting vehicles and the same target vehicle. For any assisting vehicle h , first initializing the behavior space of h (line 2), and putting all possible combinations into the set $A(h)$. Obviously, the number of combinations in $A(h)$ is 2^n . AP maintains an assisting vehicle sequence $List(i)$ which has been selected to provide the assisted download service for any target vehicle i within BA. Once an assisting vehicle is selected for the target vehicle i , it is added to $List(i)$ (the update operation of $List$ is performed in line 9 of the algorithm fragment in Figure 9). For any target vehicle i (line 3) in the target vehicle set C , this algorithm traverses every selected vehicle in $List(i)$ (line 4). If the current assisting vehicle h conflicts with the selected assisting vehicle

j (line 5), deleting all combinations with i in $A(h)$ (line 6). Repeating this operation on lines 8–9 until all combinations are checked. Line 10 returns the optimized $A(h)$ set.

The above two optimization algorithms are used to simplify the behavior space. The impossible vehicle selection behaviors are deleted from the possible behaviors of each assisting vehicle to reduce the computing complexity of the system. The follow-up experiments show that when the target traffic density $\lambda = 10$, the complexity of the algorithm can be reduced by 2 to 3 orders of magnitude.

(3) Balanced the number of downloads

In addition to maximizing the total number of downloads for users, another principle that the vehicle selection scheme needs to follow is to make every user who requests the download have a balanced access to the assisted download services. For example, in a period of time, AP can provide k services to n target vehicles through m assisting vehicles (an assisting vehicle can provide services for one or more target vehicles, so $k = m$). So, the ideal balanced download means that every target vehicle receives k/n times assisted download services.

Figure 7 is an example for a balanced download. The vehicles located under the double dotted line denotes the target vehicles, running from left to right, marked with Arabic numerals. The vehicles located above the double dotted line denote the assisting vehicles, running from right to left, marked with English lowercase letters. It is assumed that the assisting vehicles $a \sim e$ have selected the corresponding target vehicles to provide the assisted download services. As shown in the Table 1 in Figure 7, for example, item 1 indicates that it has been determined that the assisting vehicle a will provide the assisted download services for target vehicles 1, 2, and 4. The behavior space of the current assisting vehicle f is simplified to the 6 items in the

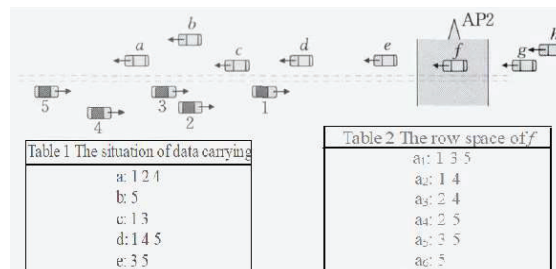


Figure 7 Analysis of balanced download cases.

Table 2 in Figure 7 through the two algorithms proposed in the previous section. Now if the assisting vehicles $a \sim e$ have selected the corresponding target vehicles to provide the assisted download services, what kind of decision should the AP make on f to make the target vehicles 1~5 closer to the balanced services?

In Figure 7, before selecting which kind of vehicle selection behavior of f , there are 5 assisting vehicles providing 11 downloads for the 5 target vehicles. Then the balance means that every target vehicle gets 11/5 downloads on average (number of downloads/number of target vehicles). At this time, the target vehicles 1~5 got 3, 1, 2, 2, and 3 downloads respectively. How far is the current situation from the balanced? Which behavior of f can bring the result closer to the balanced value? For these problems, the proposed scheme creatively uses multidimensional Euclidean Distance and Manhattan distance as the metric. By solving Euclidean Distance or Manhattan Distance from each decision option in the behavior space to the balanced service as the basis for the decision-making of AP, and using the balanced download scheme to define the probability P of the recursive formulas (7) and (8), the formulas can be solved.

(1) Euclidean Distance

Euclidean Distance (ED) is the shortest distance between two points. If the number of services is regarded as a vector, the average number of services obtained by every vehicle is the origin, and the actual number of obtained services is the straight-line distance to the origin, that is, Euclidean distance. Then the probability that the option leading to the shortest distance to the origin among all options of the current target vehicle is selected should be the greatest.

Assuming that there are n target vehicles in the target vehicle set $C = \{c_1, c_2, \dots, c_n\}$, then the number of the existing services of every target vehicle at time t corresponds to formula (11),

$$U(t) = \{u_1^t, u_2^t, \dots, u_n^t\} \quad (11)$$

then ED in the current situation can be denoted formula (12).

$$\begin{aligned} d_e(U(t)) &= \sqrt{\left[\frac{\sum_{i=1}^n u_i^t}{n} - u_1^t\right]^2 + \dots + \left[\frac{\sum_{i=1}^n u_i^t}{n} - u_n^t\right]^2} \\ &= \sqrt{\sum \left[\frac{\sum_{i=1}^n u_i^t}{n} - u_i^t\right]^2} \end{aligned} \quad (12)$$

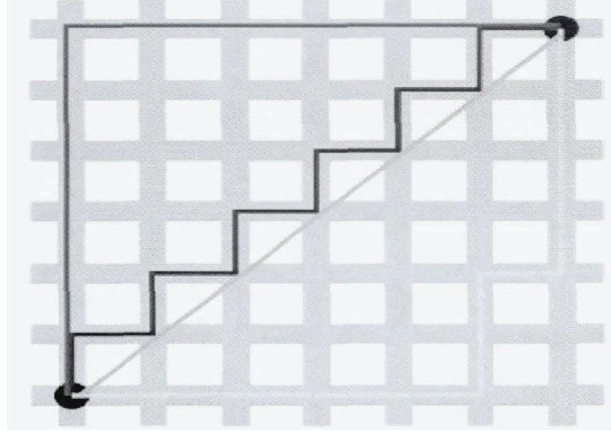


Figure 8 Euclidean distance and manhattan distance.

At time $t + 1$, for the behavior space $A(h)$ of the assisting vehicle h , the closer to balance the option can minimize $d_e(t + 1)$ is also, the higher the probability of AP selecting the behavior. It may appear that the $d_e(t + 1)$ s calculated by all options are greater than $d_e(t)$, which denotes that selecting any option will make the system more unbalanced than the current one. However, in order to increase the total amount of user download, AP should select the behavior with the smallest $d_e(t + 1)$.

Obviously, the complexity of ED between two points is n^2 . In the case of a large number of assisting vehicles, computing ED will increase the processing delay of system. Therefore, in the case of relatively large number of target vehicles, the proposed scheme uses another calculation method with relatively low computational complexity, which is to calculate the coordinate distance between two points—Manhattan Distance(MD).

(2) Manhattan Distance

As shown in Figure 8, the straight-line distance between two points is ED discussed above, and the other three distances are all coordinate distances, namely MD. The computing formula is as follow.

$$\begin{aligned}
 d_m(U(t)) &= \left| \frac{\sum_{i=1}^n u_i^t}{n} - u_1^t \right| + \dots + \left| \frac{\sum_{i=1}^n u_i^t}{n} - u_n^t \right| \\
 &= \sum_{i=1}^n \left| \frac{\sum_{i=1}^n u_i^t}{n} - u_i^t \right| \tag{13}
 \end{aligned}$$

Like ED, at time $t + 1$, for the behavior space $A(h)$ of the assisting vehicle h , the closer to balance the option being able to minimize $d_m(t+1)$ is also, the higher the probability of AP selecting the behavior. Compared with ED, the computing complexity of MD is n , which makes the proposed scheme with MD more suitable for the situations where there are relatively many target vehicles. However, the computing accuracy of MD is not as high as ED. The experiments compare the total number of user downloads using the above two distance algorithms as the basis for the decision-making.

(3) Model solution

The significance of computing ED and MD (hereinafter collectively referred to as the balanced distance, denoted by d) is that how does the current assisting vehicle provide the download services to the target vehicles to make the target vehicle in the system closer to obtaining evenly services on the premise that the previous assisting vehicle has determined to provide the assisted download services to which target vehicles. However, besides of the balanced services obtained by users, another goal is to maximize the system benefit proposed by the formula (5). In order to make the decision for selecting vehicle to reflect the balanced service and the maximum benefit of the system at the same time, the proposed scheme proposes to use the balanced distance as the probability of solving the formula (7) and (8).

Assuming that there are m kinds of behaviors in $A(h) = \{a_1, a_2, \dots, a_m\}$, the corresponding service times are respectively $U_1(t), U_2(t), \dots, U_m(t)$, and the corresponding balanced distances are $d(U_1(t)), d(U_2(t)), \dots, d(U_m(t))$. The larger the distance, the more deviation from the balanced value when selecting this kind of behavior so that the probability selecting the corresponding behavior is smaller. So, the proposed scheme uses an inverse proportional function to define the probability of this option being selected. The probability P of the formula (7) and (8) are defined as follows.

$$P(s_{t+1}|s_t, a_t^i) = \frac{1}{d(U_i(t))}{\sum_{i=0}^m \frac{1}{d(U_i(t))}} \quad (14)$$

Figure 9 shows a fragment of the recursive solution for the model. The behavior space of each $A(h)$ has been simplified relying on the simplification algorithm proposed above. As a global variable, π is the final solution sequence. For each vehicle selection behavior in $A(h)$ (line 3), if there is a next vehicle (line 4), calculating the benefit from the behavior of the next vehicle (line 4) according to the probability of the formula (13) (Line 5). If

```

1.  getMax(h) begin
2.  J ← 0
3.  for each i in A(h)
4.    if A(h+1) is ture
5.      v = P(st+1|st, at)getMax(h+1);
6.      if (J < J(i) + v)
7.        J ← J(i) + v;
8.        π(t+1) ← h+1, i;
9.        update List
10.     end
11.    else
12.      if (J < R(st, at))
13.        J ← R(st, at);
14.        π(t) ← h, i;
15.      end
16.    end
17.  end
18.  return J;
19.  end

```

Figure 9 Recursive equation solving algorithm.

the sum of the benefit and the obtained benefit from the current decision i is greater than the existing maximum value, then recording it in π (line 8) and updating the list $List$ (line 9). Here, $List$ is a global variable. And if the preorder decisions are different, it will affect $List$. If the current vehicle is the m th assisting vehicle, finding the maximum benefit and assigning it to J (line 13), updating π (line 14), finally returning to J (line 19).

By the above-mentioned solving process, it has obtained the global optimum solution for the selection scheme of assisting vehicle with a step length of m . Within BA with 16 km, when the traffic flow $\lambda = 10$, the target vehicle may encounter about 60 assisting vehicles. Obviously, it is unrealistic to take the global optimum for 60 assisting vehicles ($m = 60$). The main reasons are: (1) The vehicle speed cannot be kept constant even in a highway where the speed change rate is very low. The more vehicles involved in the computing and the farther away from AP, the less accurate the estimated speed of the vehicles will be. (2) The more vehicles involved in the computing, the more complicated the computing.

Based on the above analysis, the proposed scheme makes the number of assisting vehicles involved in the computing $m < 8$. So, the solution obtained from the proposed scheme is not a true global optimum solution, but an approximate optimal solution. When the number of target vehicles $n < 5$, the proposed scheme uses ED as the probability of vehicle selection; when

$n > 5$, the proposed scheme uses MD as the probability of vehicle selection. This method reduces the processing delay of the system.

3 Performance Evaluation

In this section, the performance indicators such as the amount of download from the proposed scheme are evaluated by simulation experiments. A part of the vehicle running data use the test data of Electric Vehicle R&D Center (Shanghai) of Chinese Academy of Sciences in Shanghai section of highway G15 from September to December 2010, and the other part is automatically generated by a data generator. During the process of generating data, the vehicle speed during the data generating period is took randomly between 90 km/h and 150 km/h, which is in line with the actual situation of the highways. Here, the probability p of the vehicle speed change(the change range being from 90 km/h to 150 km/h) conforms to Lognormal distribution, and the traffic flow conforms to Poisson distribution [13]. In the experiments, the target vehicle speed is assumed to be 90 km/h. The simulation tool is J-DSRelay that a highway vehicle networking simulation software based on dynamic time slot using JAVA to develop. The test result is the average of 10–20 tests.

From the reference [11], CA of AP on the highways are set to 800 m, and the distance between the two AP is 16 km, which is in line with the actual situation AP generally locating at the gas-station or service-area on the highways. The communication radius of the vehicle is set to 300 m. During the assisted download process, the time spending on the target vehicle establishing a connection with the assisting vehicle is 2 s [14]. Considering noise and coherence time, the transmission performance from the transmission rate of 6 Mbps (four-bit phase modulation, QPSK) is higher than 18 Mbps (sixteen-bit quadrature amplitude modulation, 16-QAM) and 3 Mbps (two-bit phase modulation, BPSK) [15], and the packet transmission rate of the link layer can reach about 60%. So, the transmission rate of the link layer between vehicles can reach 400 KB/s. However, considering the redundancy and retransmission mechanism of data packet header and end based on TCP/IP protocols, the proposed scheme sets the speed rate of download within AP area to 150 KB/s according to the relative speed difference, and the speed rate of assisted download of the opposite vehicle (the maximum relative speed) is set to 50 KB/s, and the speed rate of assisted download of the same direction vehicle (the minimum relative speed) is set to 200 KB/s [11].

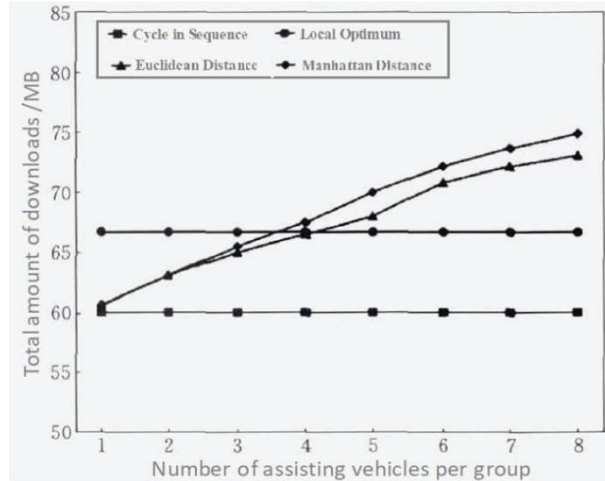


Figure 10 Comparison of the amount of download under different number of assisting vehicles.

For the different m , Figure 10 tests the total amount of user download using four vehicle selection schemes (the sequential cycle, the local optimum, ED method, and MD method). It can be seen that the sequential cycle and local optimum vehicle selection are not affected by the value of m . Since the sequential cycle is based on the number of download times of balancing the target vehicle, the total account of download is the lowest, about 60M, and while the local optimum total download is about 67M. Since adopting ED and MD method, when the proposed scheme solves the formulas, as the number of vehicles involved in the computing gradually increases, the total amount of download gradually increases. When $m = 4$, the two methods have reached the amount of download of the local optimum. When $m = 8$, the total amount of download is increased to about 75M. The total amount of download is higher than the previous two methods, and achieves the goal of balance. Figure 10 also shows that MD method can obtain more data download than ED method.

In order to be able to clearly describe the experimental results, this paper assumes that the transmission rate between vehicles is the same. As mentioned in the previous paper, the proposed scheme uses the same length of time slice in transmission so as to be able to allocate the resources, so every time the time of the assisted download services is the same. This can conclude that the amount of download for each service is the same, that is,

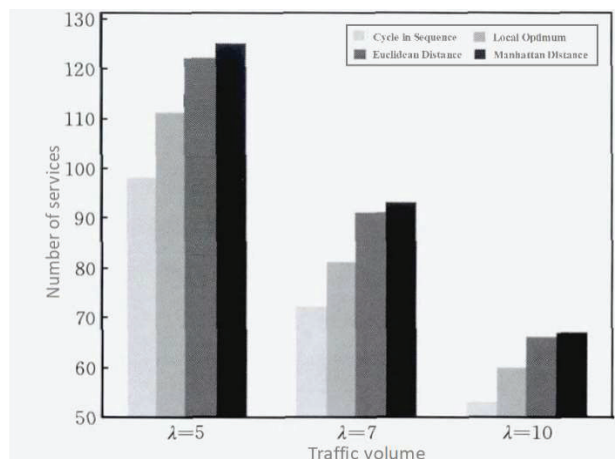


Figure 11 Comparison of 4 schemes for different traffic flows.

the system revenue = total number of services \times transmission rate \times time slice. Since the latter two are fixed, the total number of services is used as the evaluation index of the system revenue in the experiments.

Figure 11 shows the number of services under the different traffic flows. When $\lambda = 5$, the number of service times of the four schemes is 98 times, 111 times, 122 times and 125 times respectively. Compared with the vehicle selection scheme with cycle in sequence, the proposed scheme increases the number of services by about 25%. When $\lambda = 7$ and $\lambda = 10$, the number of service times obtained by the four vehicle selection schemes all have different degrees of decline. Compared with the sequential cycle and the local optimum vehicle schemes, the increment of the proposed scheme also decreases significantly, which means that the traffic flow is greater, the more obvious the advantage of the proposed scheme.

Figure 12 shows that the amount of download within a download area and a blind area (running time about 350s) using 4 assisted download schemes—MobTorrent [4], DSRelay [13], DCPD [16], and the proposed scheme. The first 30s in CA of AP, the amount of download of the four schemes is the same. There is no the assisting vehicle running in the opposite direction for about 30s to 150s, and the amount of download does not increase. After 150s, with the increase in the number of assisting vehicles, the total amount of download of the target vehicle continues to increase. It can be seen from the results that after using the proposed scheme to plan BA, the total amount of download is significantly higher than the other three schemes after 220s. The

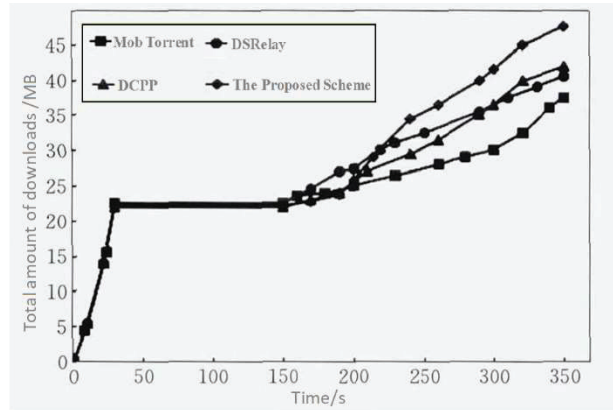


Figure 12 Comparison of the amount of download of 4 assisted download schemes.

Table 1 Comparison of the number of provided services using 5 methods

The Target Vehicle	The Sequential Cycle	The Local Optimum	The Approximately Optimum	ED	MD
c_1	20	32	33	27	29
c_2	20	18	27	23	25
c_3	20	12	20	25	22
c_4	19	21	31	22	28
c_5	19	28	23	24	21
Total	98	111	134	122	125

total amount of download is 47.2M before reaching CA of next AP, which is about 20% higher than the other three schemes (39.9M, 38.6M, 36.2M respectively).

Table 1 compares the number of services provided to 5 target vehicles by the sequential cycle, the local optimum, the proposed scheme without the balanced method, and the proposed schemes with introducing ED method and MD method respectively. For the vehicle selection scheme with cycle in sequence, the principle of vehicle selection is to evenly allocate the download services for each assisting vehicle. So, 98 services provided by the assisting vehicle are evenly allocated to 5 target vehicles in this scheme. And for the vehicle selection scheme with the local optimum, the number of provided services is 111. Compared with the sequential cycle, the number of services is significantly increased. However, c_1 with obtaining the most services (32 assisted download services) obtain 20 more services than c_3 with obtaining the least services (12 assisted downloads). Obviously, this vehicle selection

scheme can lead to the unbalanced service allocation. For the proposed scheme, which don't consider the requirement for the balanced download, it can increase the number of downloads to 134, an increase of about 20%, but there is still a large difference between the maximum number and the minimum number. The data in the last two columns are the number of services obtained by each target vehicle using the proposed scheme, which's vehicle selection is achieved based on ED method and MD method respectively. The maximum and minimum number differ by 5 and 8, respectively, which are relatively balanced. But the number of obtaining services has decreased slightly.

By the above experiments, the following conclusions can be drawn:

- (1) Compared with the sequential cycle and the local optimum selection schemes, using the proposed scheme can increase the total amount of user download by more than 20%.
- (2) Using ED and MD methods to balance the number of downloads, although the total number of user downloads will decrease slightly, the number of assisted download services obtained by each target vehicle tends to be more balanced. This scheme promotes the fairness of the system operation.

4 Conclusion

In order to effectively use the time-space resources within the blind area of AP on highway, to maximize the total number of user downloads on the premise of ensuring the fairness of the system, this paper proposes a assisted download and vehicle selection scheme for VANET. The proposed scheme innovatively proposes to use a two-dimensional matrix to define the time-space resources within the blind area and the vehicle selection behavior, and utilize the communication features of VANET to simplify the vehicle selection behavior space so as to reduce the computing complexity. The optimization model based on MDP is established to solve the problem of time-space resource allocation within the blind area. At the same time, the proposed scheme proposes to computing ED and MD as the basis of vehicle selection to achieve the balanced services providing to the target vehicle, effectively increasing the total amount of user download.

The proposed scheme utilizes the features of VANET on the highway and proposes an approximate global optimum resource allocation algorithm. Compared with the sequential cycle and local optimum vehicle selection

schemes in the existing researches, the proposed scheme not only increase the total amount of user download to a certain extent, but also keep the fairness of the system.

References

- [1] Nandan A, Das S, Pau G, et al. Cooperative downloading in vehicular Ad Hoc Wireless networks//Proceedings of the International Conference on Wireless on Demand Network Systems and Services. St. Moritz, Switzerland, 2005: 32–41.
- [2] Chen Bin-Bin, Chan Mun-Choon. MobTorrent: A Framework for mobile Internet access from vehicles//Proceedings of the IEEE INFOCOM. Barcelona, Spain, 2009: 1–9.
- [3] Wu Y, Zhu Y, Li B. Infrastructure-assisted routing in vehicular networks//Proceedings of the IEEE INFOCOM. Orlando, USA, 2012: 48–57.
- [4] Hung C C, Chan H, Hsiao E. Mobility pattern aware routing for heterogeneous vehicular networks// Proceedings of the IEEE WCNC. Las Vegas, USA, 2007: 244–249.
- [5] Ng S C, Zhang Wu-Xiong, Yang Yang. Analysis of access and connectivity probabilities of the IEEE WCNC. Sydney, Australia, 2010: 434–440.
- [6] Trullols-Cruces O, Fiore M, Barcelo-Ordinas J. Cooperative download in vehicular environments. *IEEE Transactions on Mobile Computing*, 2012, 11(1): 663–678.
- [7] Zheng Z, Sinha P, Kumar S. Alpha coverage: Bounding the interconnection gap for vehicular Internet access//Proceedings of the IEEE INFOCOM 2009. Rio de Janeiro, Brasil, 2009: 721–730.
- [8] Zheng Z, Sinha P, Kumar S. Maximizing the contact opportunity for vehicular Internet access// Proceedings of the IEEE INFOCOM. San Diego, USA, 2010: 481–489.
- [9] Su Z, Ren P, Xu R, Katto J. A novel algorithm to control contents selectively for vehicular communication networks//Proceeding of the IEEE Vehicular Technology Conference Fall. Ottawa, Canada, 2010: 1011–1017.
- [10] Lin Chuang, Wan Jian-Xiong Xiang Xu-Dong, et al. Dynamic optimization in computer systems and computer networks: Models, solution, and applications. *Chinese Journal of Computer*, 2012, 35(7): 1339–1357.

- [11] Gerlough D L. Poisson and Traffic: Use of Poisson Distribution in Highway Traffic. Eno Foundation for Highway Traffic Control, 1955.
- [12] Liu Jian-Hang, Bi Jing-Ping, Xu Peng, et al. A compensation model of cooperative downloading improving system throughput. Chinese Journal of Computers, 2012, 35(7): 1390–1398.
- [13] Liu Jian-Hang, Bi Jing-Ping, Bian Yong-Chao, et al. DSRelay: A scheme of cooperative downloading based on dynamic slot//Proceedings of the IEEE ICC, Ottawa, Canada, 2012: 621–627.
- [14] Bychkovsky V, Hull B, Miu A K, et al. A measurement study of vehicular Internet access using in situ Wi-Fi networks//Proceedings of the International Conference on ACM MobiCom. Los Angeles, USA, 2006: 891–897.
- [15] Bai F, Stancil S, Krishnan H. Toward understanding characteristics of Dedicated Short Range Communications(DSRC) from a perspective of vehicular network engineers//Proceedings of ACM MobiCom, Chicago, USA, 2010: 471–480.
- [16] Altman E, Sassatelli L, De Pellegrini F. Dynamic control of coding for progressive packet arrivals in DTNs. IEEE Transactions on Wireless Communications, 2013, 12(2): 32–51.

Biographies



Qibin Zhou, received her M.Sc. degree from Huazhong University of Science and Technology in 2013. She is the deputy director of the Information Office and the vice president of the college of International Education in Shanghai Jianqiao University now. Her present research interests include network security, data analysis, smart campus construction, etc.



Qinggang Su, received the B.Sc. degree in Computer Science from Anhui University of Technology in 2002, and got the M.Sc. degree in Communication Engineering in Shanghai Jiao Tong University, and is studying for Ph.D. degree at East China Normal University. He became a faculty member in the school of electronic information, Shanghai Dianji University China from 2002, and he is the vice dean of Chinesisch-Deutsche Kolleg für Intelligente Produktion of Shanghai Dianji University now. He is a member of China Computer Federation (CCF), and his research is currently focused on wireless networks, 5G application and smart manufacturing.



Peng Xiong, received the B.Sc. degree and M.Sc. degree in Electrical Engineering from Nanchang University China in 1998 and 2004, respectively, and Ph.D. degree in Computer science and technology from East China Normal University China in 2009. From 2010 on, he is a faculty member in the school of electronic information, Shanghai Dianji University China. He is a member of China Computer Federation (CCF), and his research is currently focused on network secure, wireless networks, cloud computing and big data etc.

