Nearly Optimal Probabilistic Coverage for Roadside Advertisement Dissemination in Urban VANETs

Yawei Hu, Mingjun Xiao, Liusheng Huang, Ruhong Cheng and Hualin Mao
School of Computer Science and Technology / Suzhou Institute for Advanced Study,
University of Science and Technology of China, Hefei, P. R. China

Abstract—Advertisement disseminations based on Roadside Access Points (RAPs) in vehicular ad-hoc networks (VANETs) attract lots of attentions and have a promising prospect. In this paper, we focus on a roadside advertisement dissemination, including three basic elements: RAP Service Provider (RSP), mobile vehicles and shops. The RSP has deployed many RAPs at different locations in a city. A shop wants to rent some RAPs, which can disseminate advertisements to vehicles with some probabilities. Then, it tries to select the minimal number of RAPs to finish the advertisement dissemination, in order to save the expenses. Meanwhile, the selected RAPs need to ensure that each vehicle’s probability of receiving advertisement successfully is not less than a threshold. We prove that this RAP selection problem is NP-hard. In order to solve this problem, we propose a greedy approximation algorithm, and give the corresponding approximation ratio. Further, we conduct extensive simulations on real world data sets to prove the good performance of this algorithm.

Index Terms—Vehicular Ad-hoc Network; Advertisement Dissemination; Roadside Access Points Selection;

I. INTRODUCTION

Vehicular ad-hoc networks (VANETs) as a new paradigm of mobile ad-hoc networks, have attracted attentions of many researchers. With the development of intelligent transportation system, some cities have deployed many Roadside Access Points (RAPs) at different locations, such as bus stations, taxi pickup points and intersections. These RAPs have good capacity of storage and communication. When a vehicle enters the communication range of a RAP, the RAP can deliver the information stored in its memory to the vehicle via WiFi. By using these RAPs, shops can disseminate large-size multimedia advertisements to passengers in vehicles at a low cost. Due to this advantage, much research ([1], [2], [3]) has studied the roadside advertisement dissemination based on RAPs in VANETs.

Consider a scenario of roadside advertisement dissemination. There is a RAP Service Provider (RSP) who has deployed many RAPs at different locations in a city. If a shop hopes to publish some advertisements, it can rent parts of the RAPs from the RSP to disseminate these advertisements. On the one hand, they need to select as less RAPs as possible for this shop to conduct the advertisement dissemination, in order to save the shop’s expenses. On the other hand, they also need to let these RAPs cover adequate vehicles, so that the advertisements can be received by as many passengers as possible. Moreover, as the communication connections between RAPs and mobile vehicles are unstable and intermittent, the selected RAPs also need to ensure that the probability of each vehicle receiving the advertisements successfully is not less than a threshold.

Let’s take Fig. 1 as the example. There are 6 RAPs \( a_A, \cdots, a_E \) and 3 mobile vehicles \( (v_1, v_2, v_3) \). Suppose that a shop hopes to disseminate its advertisements to the three vehicles, and expects that the probability of each vehicle receiving the advertisements successfully is not less than the threshold 0.7. As an example, we assume that the shop selects \( \{a_B, a_C, a_D, a_E\} \) to perform the advertisement dissemination. Although each vehicle cannot receive the advertisements successfully from each single RAP with the probability no less than 0.7, their joint probability exceeds the threshold. For instance, \( v_1 \)’s probabilities of receiving the advertisements successfully from \( a_B \) and \( a_D \) are 0.6 and 0.5, respectively. Their joint probability is \( 1 - (1 - 0.6) \cdot (1 - 0.5) = 0.8 \) beyond the threshold 0.7.

For the schemes of data dissemination in VANETs, most of them pay attentions to the delivery ratio or delay of messages. However, our roadside advertisement dissemination focuses on how to select the minimal number of RAPs to conduct the dissemination. Unlike other schemes of advertisement dissemination, our scheme can ensure that each vehicle’s probability of receiving advertisement successfully is not less than a predefined threshold. For example, in [2], the authors select several seed vehicles to disseminate advertisements to other vehicles, and the seed vehicle selection is based on the degree centrality [4] of each vehicle. Nevertheless, the
solution cannot guarantee that each vehicle’s probability of receiving advertisement successfully is large enough. From another perspective, the advertisement dissemination can be regarded as the problem of coverage. For example, in [2], the dissemination is achieved by selecting some seed vehicles to cover as many vehicles as possible. Their “cover” means that there is at least one social contact (the contact strength beyond a threshold) between a single seed vehicle and each vehicle. In contrast, our problem is to ensure the joint dissemination probability from all selected RAPs beyond the threshold.

In this paper, we first describe the moving pattern of the vehicles. We define the route of a vehicle to model its mobility, and the route is a set of RAPs which this vehicle often passes. When a vehicle moves on its route, the vehicle can receive advertisements successfully from these passed RAPs with some probabilities. Then we define the RAP selection problem for roadside advertisement dissemination and prove the NP-hardness of this problem. The problem is to select the minimal number of RAPs to cover all the routes of the vehicles and ensure that each vehicle can receive advertisements successfully from the RAPs with a probability no less than a threshold. Finally, we propose an greedy approximate algorithm to solve the RAP selection problem, analyze the complexity and approximation ratio of this algorithm and conduct extensive simulations to prove its good performance.

We highlight our main contributions as follows:

1) We propose a scheme for roadside advertisement dissemination based on probabilistic coverage of some selected RAPs.

2) We introduce a new optimization problem, which is the RAP selection problem for roadside advertisement dissemination. We prove that this problem is NP-hard.

3) We propose a greedy approximation algorithm to solve the RAP selection problem, give the corresponding approximation ratio of this algorithm and conduct extensive simulations to prove its good performance.

The remainder of this paper is organized as follows. Section II presents the related work. In Section III, the network model, the definition of RAP selection problem and the proof of the NP-hardness of this problem are described. We propose an greedy algorithm, analyze its complexity and approximation ratio in section IV. In Section V, the evaluation of this algorithm is showed, followed by the conclusion of this paper in Section VI.

II. RELATED WORK

The scenario of advertisement dissemination in VANETs is different from the online advertisements ([5], [6], [7]). In VANETs, J. Qin et al. [2] considered to select seed vehicles to diffuse advertisements to others. They analyzed the sociability of the vehicular network and proved the dynamic and temporal correlations of sociality. Based on the analysis, they proposed a greedy scheme to solve it. Z. Li et al. [3] proposed an advertisement diffusion scheme with an incentive-centered architecture. The architecture encourages the advertisement providers to trade off the effect and cost of their advertisements messages, aiming to avoid unnecessary distractions to drivers and message storms in VANETs. In [1], H. Zheng et al. designed a system, which can disseminate advertisements via some placed RAPs. When a driver receives a shop’s advertisements from RAPs, he or she may detour to the shop and the detour probability depends on the detour distance. The RAP placement needs to balance the tradeoff between the traffic density and the detour probability. S.-B. Lee et al. [8] proposed a secure incentive framework to avoid the noncooperative behavior of selfish or malicious vehicle nodes in the advertisement dissemination. Unlike these schemes, our paper uses the RAPs to disseminate advertisements to the vehicles passing by.

Our roadside advertisement dissemination is relevant to the data dissemination based on roadside units (RSUs) in VANETs. For example, in [9], J. Jeong et al. considered to forward data from stationary APs to moving vehicles. The forwarding scheme took into account the AP’s location and the destination of the vehicle’s trajectory, and selected a target point as packet-and-vehicle-rendezvous-point. M. Sardari et al. in [10] proposed a message dissemination paradigm, in which each RSU encoded a huge message into \( k \) data packets and forwarded them to vehicles, then the vehicles can decode a specific RSU’s message by collecting sufficient packets. K. Liu et al. [11] focused on accessing information stored at RSUs and paid attention to the channel division. Comparing with these papers, we consider how to select the minimal RAPs to finish the advertisement dissemination from RAPs to vehicles.

Our problem, i.e., the RAP selection, is related to the deployment of RSUs in VANETs. In [12], Wu et al. studied the RSU placement problem for vehicular networks in a highway-like scenario. Their placement strategy maximized the aggregate throughput in the network, taking into account the impact of wireless interference, vehicle population distribution, and vehicle speeds in the formulation. In [13], B. Aslam et al. focused on the placement problem in the scenario of urban vehicular network environment. They proposed two methods aiming to minimize the reporting time of event for RSUs, with incorporating the density and speed of vehicle, as well as the occurrence likelihood of an event in urban. By comparison, our problem is based on the joint probabilistic coverage of RAPs.

Therefore, our roadside advertisement dissemination is different from other studies of advertisement or data dissemination, and cannot be solved directly by existing solutions.

III. PROBLEM DEFINITION

In this section, we set up the network model, define the RAP selection problem for roadside advertisement dissemination and analyze the hardness of this problem.

A. Network Model

We first describe the set of \( i \) vehicles, that is, \( V = \{v_1, v_2, \cdots, v_i\} \). We assume that the vehicles have been equipped with wireless communication devices. Consequently, they can communicate with the RAPs. Then, we use the
set $A = \{a_1, a_2, \ldots, a_j\}$ to denote all the RAPs which are deployed by the RSP at different locations in a city. As we know, a vehicle often visits some locations frequently due to the sociality of the vehicle (actually, the driver or the passenger). For simplicity, we describe the route of each vehicle with a set of locations where the vehicle passes by and the RAPs are placed at. Namely, the route of vehicle $v_i$ is $R(v_i) = \{a_{i_1}, a_{i_2}, \ldots, a_{i_k}\}$ ($a_{i_k} \in A$). Here, we don’t care the order of $a_{i_1} \cdots a_{i_k}$. Next, for $v_i \in V$ and $a_j \in R(v_i)$, we use $p_j^i (0 \leq p_j^i < 1)$ to indicate the probability that $v_i$ successfully receives advertisements from $a_j$. Although some vehicle $v_i$ may pass certain RAP twice or more, we only consider the vehicle passing each RAP in $R(v_i)$ once. This assumption is reasonable. If $v_i$ passes $a_j$ $n$ times, we can use the $1 - (1 - p_j^i)^n$ to replace the $p_j^i$, and meanwhile, treat it as one-time passing. In addition, we assume that the value of $p_j^i$ can be derived from the history records.

We utilize a bipartite graph $G$ to describe the model. The vertex set of $G$ is $V \cup A$. For an arbitrary pair of $v_i$ and $a_j$, if $a_j \in R(v_i)$, we add an edge $(v_i, a_j)$ into the edge set of $G$, and attach $p_{j}^{i}$ as the weight of $(v_i, a_j)$. Let’s take Fig. 2 as an example. The vehicle node set is $V = \{v_1, v_2, \cdots, v_5\}$, the RAP node set is $A = \{a_1, a_2, a_3, a_4\}$. $v_1$’s route $R(v_1)$ is $\{a_1, a_2\}$, $v_2$’s route $R(v_2)$ is $\{a_2, a_3\}$, and so on. By the way, we denote $N(-)$ as the set of neighbor nodes of one vertex. In Fig. 2, $N(a_1) = \{v_1, v_3\}$. Moreover, the degree of a vertex is denoted as $\text{deg}(-)$ and $\text{deg}(v_i) = |R(v_i)|$, $\text{deg}(a_j) = |N(a_j)|$.

### B. Problem Definition

In our scheme of roadside advertisement dissemination, a shop hopes that the vehicles in $V$ can successfully receive its advertisements from the RAPs with a probability. Then the shop needs to select as less RAPs as possible to conduct the advertisement dissemination to these vehicles, and ensure that each vehicle’s probability of receiving advertisements successfully is not less than a threshold. We use $\tau$ to denote this threshold $(0 < \tau < 1)$. Moreover, we utilize $S$ to denote the set of all selected RAPs from $A$. Then, for a vehicle $v_i$, it may receive advertisements from each RAP in $S \cap R(v_i)$, denoted by $S(i)$. As a result, $v_i$’s probability of receiving advertisements successfully from the RAPs in $S(i)$ is $P_i|S = 1 - \prod_{a_j \in S(i)} (1 - p_j^i)$ for the RAP selection solution $S$. Especially, if $S(i) = \emptyset$, $P_i|S = 0$.

In this scheme, the problem we need to solve is how to select minimal number of RAPs and ensure that each vehicle’s probability of receiving advertisements successfully exceeds the threshold $\tau$. In summary, we illustrate this problem of RAP selection as follow.

\[
\begin{align*}
\text{minimize} & : |S| \\
\text{s.t.} & : \bigcup_{a_j \in S} N(a_j) = V; \\
& \quad P_i|S \geq \tau \text{ for } \forall v_i \in V; \quad (1) \\
& \quad P_i|S = 1 - \prod_{a_j \in S(i)} (1 - p_j^i); \\
& S \subseteq A.
\end{align*}
\]

Here, we assume that $P_i|A \geq \tau$ for $\forall v_i \in V$. If there is a vehicle $v_k$ satisfying $P_k|A < \tau$, the $v_k$ cannot receive advertisements successfully with the probability no less than $\tau$, even if all RAPs in $A$ are selected. Actually, there is no feasible solution to this case. In this paper, we only study the cases which exist feasible solutions.

### C. Problem Hardness Analysis

**Theorem 1:** The RAP selection problem for roadside advertisement dissemination is NP-hard.

**Proof:** We simplify the RAP selection problem by assuming $p_j^i = 1$ for $\forall v_i \in V$ and $\forall a_j \in A$. Hence, $P_i|S \geq \tau$ for $\forall v_i \in V$ is always true and the simplified problem is to select the minimal number of RAPs from $A$ to cover all vehicles in $V$. It is obvious that $\bigcup_{a_j \in A} N(a_j) = V$ and $N(a_j) \subseteq A$. Let’s denote $N_A = \{N(a_1), N(a_2), \cdots, N(a_j)\}$, therefore, $N_A$ is a collection of $V$’s subsets. For the RAP selection problem, if to make the selection of the RAP $a_k$ corresponds to the selection of $V$’s subset $N(a_k)$, this problem finally can be illustrated as the following (2).

\[
\begin{align*}
\text{minimize} & : |N_a| \\
\text{s.t.} & : \bigcup_{N(a_k) \subseteq N_a} N(a_k) = V; \quad (2) \\
& \quad N_a \subseteq N_A.
\end{align*}
\]

Based on [14], we can conclude that Eq (2) is the same as the problem of Set Cover. Therefore, the RAP selection problem is NP-hard. □

### IV. GREEDY ALGORITHM AND PERFORMANCE ANALYSIS

In this section, we propose a greedy algorithm to solve the RAP selection problem, and analyze the correctness, complexity and approximation ratio of this algorithm.

**A. Greedy Algorithm**

Before the algorithm, we define a utility function $f(\cdot): 2^A \rightarrow Q^+$, which is a mapping from a RAP set to a real utility value. It indicates the probabilistic coverage utility of a given RAP set, in other words, the sum of probabilities that each vehicle in $V$ successfully receives advertisements from the given set of RAPs. Concretely, the utility is defined as follows.

\[
f(S) = \theta \sum_{v_i \in V} \min\{P_i|S, \tau\} \quad (3)
\]
where \( \theta = \max \{ \frac{1}{\tau}, \frac{1}{m} \} \), \( m = m_I \cdot (1 - m_u)^m \cdot m_u = \max \{ P_i | v_i(\forall v_i)\in V \land \forall a_j \in A \} \), \( m_I = \min \{ P_i | v_i(\forall v_i)\in V \land \forall a_j \in A \} \), \( m_u = \max \{ \deg(v_i) | \forall v_i \in V \} \).

In Eq. 3 \( \min \{ P_i | S, \tau \} \) is the probabilistic coverage utility of \( S \) to the vehicle \( v_i \), i.e., the (joint) probability that \( v_i \) successfully receives advertisements from the RAPs in \( S \). Moreover, along with the increase of the probability, the utility value \( \min \{ P_i | S, \tau \} \) will become larger and larger. When the probability exceeds the threshold \( \tau \), the utility value will not change any more. \( \theta \) is a parameter that we defined for the approximation ratio analysis in the following part.

Let \( S \) be the selected RAP set. When a RAP \( a_k \) is added to \( S \), some vehicles which haven’t been covered will be covered by \( a_k \), or some vehicles’ probabilities of receiving advertisements successfully will increase. Both of them result in the growth of the utility. Then, the basic idea of our algorithm is to select the RAP \( a_k \) which maximizes the growth of the utility, and add it to the set \( S \) in each iteration. The concrete algorithm is shown in Algorithm 1 in which \( S \) is initialized to be \( \emptyset \). In each iteration, the RAP \( a_k \) with the maximum \( f(S + \{ a_k \}) \) is selected and added into \( S \). The algorithm terminates when \( f(S) = \theta \cdot \tau \cdot |V| \).

**Algorithm 1** Greedy Algorithm for RSP Selection

1: Input: \( V, A, S, p_i^j, \tau \)
2: Start:
3: \( S = \emptyset \)
4: While \( f(S) < \theta \cdot \tau \cdot |V| \) do
5: choose the RAP \( a_k \in A \setminus S \) to maximize \( f(S + \{ a_k \}) \);
6: \( S = S \cup \{ a_k \} \);
7: End
8: Output: \( S \)

We use the example in Fig. 2 to illustrate Algorithm 1 in which \( \tau = 0.6 \), \( \theta = 3125 \) and denote \( f_i = \min \{ P_i | S, \tau \} \). Fig. 3 shows the corresponding results.

1) First iteration: \( f(\{ a_1 \}) = 2812.5 \), \( f(\{ a_2 \}) = 5000 \), \( f(\{ a_3 \}) = 3750 \), \( f(\{ a_4 \}) = 3750 \). Hence, we select \( a_2 \) and add it to \( S \). After this iteration, \( P_1 | S = 0.4 \), \( P_2 | S = 0.6 \), \( P_3 | S = 0.0 \), \( P_4 | S = 0.4 \), \( P_5 | S = 0.2 \). That is to say, \( v_2 \) is covered with its probability of receiving advertisements successfully no less than 0.6, \( v_3 \) isn’t covered, \( v_1, v_4, v_5 \) are covered with their probabilities of receiving advertisements successfully less than 0.6.

2) Second iteration: \( f(S \cup \{ a_1 \}) = 7187.5 \), \( f(S \cup \{ a_3 \}) = 6875 \), \( f(S \cup \{ a_4 \}) = 7812.5 \). Therefore, \( a_4 \) is selected and added into \( S \). After this iteration, both \( v_4 \) and \( v_5 \) can receive advertisements successfully with the probability no less than 0.6.

3) Third iteration: \( f(S \cup \{ a_1 \}) = 9375 \), \( f(S \cup \{ a_3 \}) = 8750 \). Therefore, \( a_1 \) is selected and added into \( S \).

4) Fourth iteration: we compute \( f(S) = 9375 = \theta \cdot \tau \cdot |V| \).

Therefore, the algorithm terminates. As \( f_i = 0.6 \) for \( \forall v_i \in V, S = \{ a_1, a_2, a_4 \} \) is a feasible solution.

**B. Correctness, Complexity and Approximation Ratio**

We first show the correctness of Algorithm 1 by theoretical analysis. On the one hand, in Algorithm 1 only one RAP is added into \( S \) in each iteration. In the worst case, all RAPs in \( A \) are added into \( S \) after \( |A| \)-th iteration. For each vehicle, if all RAPs in its route are selected, its probability of receiving advertisements successfully is no less than \( \tau \). In this case, \( f(S) \) must be \( \theta \cdot \tau \cdot |V| \), as \( \min \{ P_i | S, \tau \} = \tau \) for \( \forall v_i \in V \). Therefore, Algorithm 1 terminates for sure. On the other hand, if \( f(S) = \theta \cdot \tau \cdot |V| \), \( \min \{ P_i | S, \tau \} \) must equal to \( \tau \) for \( \forall v_i \in V \). Therefore, each vehicle in \( V \) can receive advertisements successfully from the RAPs in \( S \) with a probability no less than \( \tau \), namely, \( S \) is a feasible solution of the RAP selection problem. In turn, if \( S \) is a feasible solution, \( f(S) \) must be \( \theta \cdot \tau \cdot |V| \) after each RAP in \( S \) is selected. In summary, our Algorithm 1 is correct.

Then, we analyze the time complexity of Algorithm 1. In Algorithm 1 the loop body will run \( |A| \) times in the worst case. During each iteration of Algorithm 1, in order to choose the maximal \( f(S + \{ a_k \}) \) for \( \forall a_k \in A \setminus S \), the algorithm needs to traverse the RAPs in \( A \setminus S \). Moreover, Algorithm 1 also need to compute \( f(S + \{ a_k \}) \), so as to compare the values of \( f(S + \{ a_k \}) \) for \( \forall a_k \in A \setminus S \). The time complexity of compute \( f(S) \) is \( O(|V| + |A|^3) \), according to Eq. 3. In sum, the time complexity of Algorithm 1 is \( O(|V| + |A|^3) \).

Finally, we give the approximation ratio of Algorithm 1 via the following Theorem 2. 

**Theorem 2:** Algorithm 1 produces an approximation solution with a ratio of \( 1 + \ln(\frac{|V|}{\theta \cdot \tau \cdot |V|}) \) from the optimal, where \( opt \) is the numbers of selected RAPs produced by the optimal solution.

**Proof:** As we all know, submodular function \( f(\cdot) : 2^E \rightarrow R \) relates closely to greedy algorithm. Based on submodular function, there exists a problem of Minimum Submodular Cover with Submodular Cost (MSC/SC) [15]. Given a universal set \( E \), a subset \( X \) of \( E \), and two increasing submodular functions \( f(\cdot) \) and \( c(\cdot) \) on \( 2^E \) with \( f(\emptyset) = c(\emptyset) = 0 \), the problem of MSC/SC is the minimization problem \( \min \{ c(X) | f(X) = f(E), X \subseteq E \} \), where \( c(X) \) is the cost of the subset \( X \). In this paper, for a RAP set \( S \subseteq A \) and the utility function \( f(\cdot) \), if \( f(S) = \theta \cdot \tau \cdot |V| \), \( S \) is a feasible solution of the RAP selection problem and \( f(S) = f(A) \). If we denote \( c(S) = |S| \), the RAP selection problem can be defined as \( \min |c(S)| f(S) = f(A), S \subseteq A \). That is to say, this problem is a kind of MSC/SC problem. Therefore, we can prove the submodularity of the utility function \( f(\cdot) \) and the function \( c(S) = |S| \), and analyze the approximation ratio of Algorithm 1 based on some studies on MSC/SC problem. The detail proofs are showed in appendix.

**V. Evaluation**

We conduct extensive simulations to evaluate the performance of our proposed algorithm. The compared algorithms,
the traces that we used, the simulation settings, and the results are presented as follows.

A. Compared Algorithms, Metric and Trace

In order to evaluate the performance of our proposed algorithm, some compared algorithms need to be introduced. As there’s no similar problems that have been discussed before, we design three alternative schemes as the compared algorithms. Differing from our greedy algorithm, which selects the RAP $a_k$ from $A \setminus S$ to maximize $f(S + \{a_k\})$ in each iteration until all of the vehicles are covered with their probabilities of receiving advertisements successfully no less than the threshold, the three compared algorithms depend on other strategies for RAP selection in each iteration.

1) Random Selection (RS): randomly selects a RAP in each iteration until each vehicle in $V$ is covered with its probability of receiving advertisements successfully no less than the threshold;

2) Maximum Degree First (MDF): in each iteration, selects the RAP $a_k$ with maximum degree, i.e., the RAP with maximum $\text{deg}(a_k)$ for $\forall a_k \in A \setminus S$. In [2], the authors use degree centrality to select seed vehicles. Based on degree centrality, we design the scheme of MDF;

3) Maximum Probability First (MPF): selects the RAP $a_k$ with the maximum sum of weights of the edges between the RAP node $a_k$ and the vehicle nodes in $N(a_k)$ in each iteration, i.e., selects the $a_k$ with maximum $\sum_{v \in N(a_k)} p^k_v$.

To compare the different results, we use the metric Selection Ratio that refers to the ratio of the number of the selected RAPs to the total number of all RAPs. Obviously, the value of selection ratio is between 0 and 1, and a solution with a smaller selection ratio is better.

We evaluate the metric of our greedy approximation algorithm (GAA) and the other three schemes, using the real data of bus systems in Hefei and Shanghai. The data of Hefei includes 125 bus lines with 898 bus stations covered, and the data of Shanghai includes 503 bus lines and 3850 stations. We regard each bus line as a mobile vehicle, each station as a location where deployed a RAP. Therefore, the route of a vehicle is the set of the stations covered by corresponding bus line. Further, the probabilities that each vehicle receives advertisements successfully from the RAPs in its route are generated randomly.

B. General Comparison

We first compare the performances of the four schemes with different threshold $\tau$, and $\theta$ is set to 10 as it has no effect on the selection results. The detail results are showed in Fig.4.

As RS scheme is a random algorithm, the value of RS in Fig.4 is an average of 10 running results.

In Fig.4 we can see that GAA achieves the best performance and much exceeds other three schemes, with different $\tau = \{0.1, \cdots, 0.9\}$. MPF has a close performance to RS when $\tau$ is small in Fig.4(a), and RS becomes worse extremely when $\tau$ increases. The trend is more obvious in Fig.4(b). MDF is the worst scheme under on data sets of Hefei and Shanghai. The reason of the bad performance of MDF is that, there always exists several “unpopular” bus lines in a city and most stations passed by these unpopular lines are also unpopular, namely, passed by few lines. As MDF selects the station (i.e., RAP) with maximum degree in each iteration, these unpopular stations will be selected to cover those unpopular bus lines (i.e., mobile vehicles) when most of other stations have been selected. It is because of this reason, MPF that reduces the effect of degree in each selection, RS which selects RAP randomly and independently of degree, achieve better performances. For GAA, its selection considers the coverage utility to all vehicles in each iteration, therefore, GAA achieves the best performance.

We also conclude from Fig.4 that the selection ratio of GAA doesn’t rise severely as the threshold $\tau$ increases. Therefore, an appropriate and larger value of $\tau$ can be chose without worry about the consequent bad performance.

C. Scalability of Algorithm

To get the scalability of GAA, we compare the four schemes when some new mobile vehicles are added. In simulations, we first run the four algorithms with $k$ randomly selected vehicles and the RAPs in the routes of these vehicles, then we continually add $k$ new vehicles selected randomly from the rest of all and compare the performances of the four schemes. Moreover, $\tau$ is set to 0.2, 0.6 and 0.9, respectively. When simulating on the data of Hefei, $k$ is set to 10, and
In this paper, we focus on the roadside advertisement dissemination in VANETs. Its scenario includes three basic elements: RSP, mobile vehicles and shops. We first introduce the model of the roadside advertisement dissemination. Based on the model, we propose a RAP selection problem for roadside advertisement dissemination, which is to select the minimal number of RAPs to cover all vehicles in model, and ensure that each vehicle can receive advertisements successfully with a probability no less than a threshold. Then we prove that the RAP selection problem is NP-hard by converting it to the Set Cover problem. Next, we propose a greedy approximation algorithm to solve this problem, analyze the complexity and approximation ratio of this proposed algorithm. Finally, we conduct extensive simulations to show the good performance of our greedy algorithm. The results indicate that our algorithm can select the minimum number of RAPs to finish the roadside advertisement dissemination, than other compared schemes.

VI. CONCLUSION

In this paper, we focus on the roadside advertisement dissemination in VANETs. Its scenario includes three basic elements: RSP, mobile vehicles and shops. We first introduce the model of the roadside advertisement dissemination. Based on the model, we propose a RAP selection problem for roadside advertisement dissemination, which is to select the minimal number of RAPs to cover all vehicles in model, and ensure that each vehicle can receive advertisements successfully with a probability no less than a threshold. Then we prove that the RAP selection problem is NP-hard by converting it to the Set Cover problem. Next, we propose a greedy approximation algorithm to solve this problem, analyze the complexity and approximation ratio of this proposed algorithm. Finally, we conduct extensive simulations to show the good performance of our greedy algorithm. The results indicate that our algorithm can select the minimum number of RAPs to finish the roadside advertisement dissemination, than other compared schemes.

REFERENCES

[1] Huanyang Zheng and Jie Wu. Optimizing roadside advertisement dissemination in vehicular cyber-physical systems. In IEEE International Conference on Distributed Computing Systems (ICDCS), 2015.
[2] Jun Qin, Hongzi Zhu, Yanmin Zhu, Li Lu, Guangtao Xue, and Minglu Li. Post: Exploiting dynamic sociality for mobile advertising in vehicular networks. In IEEE INFOCOM, 2014.
[3] Zhengming Li, Congyi Liu, and Chunxiao Chigan. On secure vanet-based ad dissemination with pragmatic cost and effect control. IEEE Transactions on Intelligent Transportation Systems, 14(1):124–135, 2013.
[4] Linton C. Freeman. Centrality in social networks conceptual clarification. Social Networks, 1(3):215–239, 1979.
[5] Anja Lambrecht and Catherine Tucker. When does retargeting work? Information specificity in online advertising. Journal of Marketing Research, 50(5):561–576, 2013.
[6] Shintaro Okazaki and Charles R. Taylor. Social media and international advertising: theoretical challenges and future directions. International Marketing Review, 30(1):56–71, 2013.
[7] Jianqing Chen and Jan Stallaert. An economic analysis of online advertising using behavioral targeting. MIS Quarterly, 38(2):429–449, 2014.
[8] Suk-Bok Lee, Joon-Sang Park, Mario Gerla, and Songwu Lu. Secure incentives for commercial ad dissemination in vehicular networks. IEEE Transactions on Vehicular Technology, 61(6):2715–728, 2012.
[9] Jaejoon Jeong, Shuo Guo, Yu Gu, Tian He, and David H.C. Du. Tsf: Trajectory-based statistical forwarding for infrastructure-to-vehicle data delivery in vehicular networks. In IEEE International Conference on Distributed Computing Systems (ICDCS), 2010.
[10] Mohsen Sardari, Faramarz Hendessi, and Faramarz Fekri. Infocast: A new paradigm for collaborative content distribution from roadside units to vehicular networks. In IEEE SECON proceedings, 2009.
[11] Kai Liu and Victor C.S. Lee. Rsu-based real-time data access in dynamic vehicular networks. International IEEE Conference on Intelligent Transportation Systems (ITSC), pages 1051–1056, 2010.
[12] Tsung-Jung Wu, Wanjiun Liao, and Chung-Ju Chang. A cost-effective placement in urban areas for vehicular networks. In IEEE Symposium on Computers and Communications (ISCC), 2012.
[13] Baber Aslam, Faisal Amjad, and Cliff C. Zou. Optimal roadside units placement in urban areas for vehicular networks. In IEEE Symposium on Computers and Communications (ISCC), 2012.
[14] Uriel Feige. A threshold of ln n for approximating set cover. Journal of the ACM, 45(4):634–652, 1998.
[15] Pengjun Wang, Dingzhu Du, Panos Pardalos, and Weili Wu. Greedy approximations for minimum submodular cover with submodular cost. Computational Optimization and Applications, 45(2):463–474, 2010.

Appendix

Detail Proofs of Theorem 2

As we all know, submodular function relates closely to greedy algorithm. Consider a set $E$ and a real function $f() : 2^E \rightarrow R$, $f(\cdot)$ is submodular if $\forall X \subseteq Y \subseteq E$ and $x \in \nabla f(Y)$, $f(X \cap \{x\}) - f(X) \geq f(Y \cap \{x\}) - f(Y)$. For example, the function $f(X) = |X|$ is submodular. Moreover, if a submodular function $f(\cdot)$ is increasing with $f(\emptyset) = 0$, then the $f(\cdot)$ is a polymatroid function. Based on polymatroid function, there exists a problem of Submodular Cover. For a set $X \subseteq E$, $X$ is a submodular cover of $(V, f(\cdot))$ if $f(X) = f(E)$. Further, given two polymatroid functions $f(\cdot)$ and $c(\cdot)$ on $2^E$, the problem of Minimum Submodular Cover with Submodular Cost (MSC/SC) is the minimization problem $\min\{c(X)\mid f(X) = f(E), X \subseteq E\}$, where $c(X)$ is the cost of set $X$.

In this paper, for a subset $S \subseteq A$ and the utility function $f(\cdot)$, if $f(S) = \theta_{\forall x \in Y}[V]$, $S$ is a feasible solution of the RAP selection problem and $f(S) = f(A)$. If we define $c(S) = |S|$, this problem can be described as $\min\{c(S)\mid f(S) = f(A), S \subseteq A\}$. That is to say, the RAP selection problem is a kind of MSC/SC. Therefore, we can analyze the approximation ratio of our algorithm based on some studies on submodular function.

First, we give Lemma 3 and Lemma 4 which prove that the function $f(\cdot)$ is a polymatroid function.

Lemma 3: $f(\emptyset) = 0$ and $f(\cdot)$ is an increasing function.

Proof:
1) If $S = \emptyset$, $P_i|S = 0$ for $\forall P_i \in V$, then $f_i = \min\{P_i|S, \tau\} = 0$, so $f(S) = 0$.
2) Giving any two sets $X$ and $Y$ and suppose $X \subseteq Y \subseteq A$,
we have \( P_1|X \leq P_1|Y \). Let’s denote \( f_i(X) = \min\{P_i|X, \tau\} \), \( f_i(Y) = \min\{P_i|Y, \tau\} \). The result of \( f_i(X) - f_i(Y) \) can be computed by dividing it into three possible cases:

1. \( P_1|X \leq \tau, P_1|Y \leq \tau, f_i(X) - f_i(Y) = P_1|X - P_1|Y \leq 0; \)
2. \( P_1|X \leq \tau, P_1|Y > \tau, f_i(X) - f_i(Y) = P_1|X - \tau \geq 0; \)
3. \( P_1|X > \tau, P_1|Y > \tau, f_i(X) - f_i(Y) = \tau - \tau = 0. \)

In sum, \( f_i(X) - f_i(Y) \leq 0 \) for \( \forall v_i \in V \). Moreover, \( \theta > 0 \), hence, \( f(X) \leq f(Y) \) when \( X \subseteq Y \subseteq A \). Therefore, \( f(\cdot) \) is an increasing function.

**Lemma 4:** \( f(\cdot) \) is a submodular function.

**Proof:** Given \( X \subseteq Y \subseteq A, \forall a_k \in A \setminus Y \), if \( \Delta a_k f_i(X) = f_i(X + \{a_k\}) - f_i(X) \geq \Delta a_k f_i(Y) = f_i(Y + \{a_k\}) - f_i(Y) \), \( f(\cdot) \) is a submodular function. In order to prove this, we compare the size of \( \Delta a_k f_i(X) = f_i(X + \{a_k\}) - f_i(X) \) and \( \Delta a_k f_i(Y) = f_i(Y + \{a_k\}) - f_i(Y) \), by dividing all possibilities into the following six cases. Note that, \( P_1|X + \{a_k\}| \geq P_1|X, P_1|Y + \{a_k\}| \geq P_1|Y, P_1|Y \geq P_1|X \) and \( P_1|Y + \{a_k\}| \geq P_1|X + \{a_k\}| \).

1. \( \tau < P_1|X, \tau < P_1|Y \). \( \Delta a_k f_i(X) - \Delta a_k f_i(Y) = (\min\{P_1|Y + \{a_k\}, \tau\} - \min\{P_1|X, \tau\}) - (\min\{P_1|Y + \{a_k\}, \tau\} - \min\{P_1|X, \tau\}) = (\tau - \tau) - (\tau - \tau) = 0; \)
2. \( P_1|X \leq \tau < P_1|X + \{a_k\}, \tau < P_1|Y \).
   \( \Delta a_k f_i(X) - \Delta a_k f_i(Y) = (\tau - P_1|X) - (\tau - P_1|Y) \geq 0; \)
3. \( P_1|X \leq \tau < P_1|X + \{a_k\}, P_1|Y \leq \tau < P_1|X + \{a_k\}) \).
   \( \Delta a_k f_i(X) - \Delta a_k f_i(Y) = (\tau - P_1|X) - (\tau - P_1|Y) \geq 0 \).
4. \( P_1|Y \leq \tau < P_1|X + \{a_k\}, \tau < P_1|Y \).
   \( \Delta a_k f_i(X) - \Delta a_k f_i(Y) = (\tau - P_1|X) - (\tau - P_1|Y) \)
5. \( P_1|Y \leq \tau < P_1|X + \{a_k\}, \tau < P_1|Y \).
   \( \Delta a_k f_i(X) - \Delta a_k f_i(Y) = (\tau - P_1|X) - (\tau - P_1|Y) \geq 0; \)
6. \( \tau < P_1|X + \{a_k\}, \tau < P_1|Y \).
   \( \Delta a_k f_i(X) - \Delta a_k f_i(Y) = (\tau - P_1|X) - (\tau - P_1|Y) \geq 0 \).

For the cases 5 and 6 above, we need to continue analyzing \( P_1|\{S + \{a_k\}\} - P_1|\{S\} \) for \( \forall v_i \in V \) and \( \forall a_k \in A \setminus S \). We denote \( \Delta a_k P_1|S = P_1|\{S + \{a_k\}\} - P_1|\{S\} \). In fact, for \( \forall v_i \). \( \Delta a_k P_1|S \) is the increment of its probability of receiving advertisement successfully when a new RAP \( a_k \) is added to \( S \). It is obvious that the \( a_k \) only influences the vehicles in \( N(a_k) \), such as showed in Fig. 6. In order to compute \( \Delta a_k P_1|S \), there exists three possible case below:

1. \( v_j \) is covered by \( S \) and \( a_k \). Hence, \( \Delta a_k P_1|S = [1 - \prod_{a_j \in S \setminus \{a_k\}} (1 - p_j)] - [1 - \prod_{a_j \in S \setminus \{a_k\}} (1 - p_j)] \) \( = \prod_{a_j \in S \setminus \{a_k\}} (1 - p_j) - (1 - p_k) \prod_{a_j \in S \setminus \{a_k\}} (1 - p_j) \) \( = p_k \prod_{a_j \in S \setminus \{a_k\}} (1 - p_j); \)
2. \( v_i \) is covered by \( S \) but isn’t covered by \( a_k \). Hence,
Based on Lemma 3 and Lemma 4, we know that the utility of a greedy algorithm, if the selected vehicle \( v_j \) that \( P_j|S_r < \tau \) after \( r \)-th iteration. Otherwise, \( P_j|S_r \geq \tau \) for all \( v_i \in V, f(S_r) \) must be \( \theta \cdot \tau \cdot |V| \) and \( S_r \) is a feasible solution, consequently, the algorithm will terminate.

3) In 1-st iteration, no matter which RAP \( a_f \) is selected, \( \Delta_{a_f} \cdot f(V) \) must be greater than 0 as \( f(\emptyset) = 0 \). After 1-st iteration, we assume \( a_k \) is selected during \( (r+2) \)-iteration, there must exist \( \Delta_{a_k} \cdot f(S_r) > 0 \). Otherwise, \( \Delta_{a_k} \cdot f(S_r) \) is maximized during \( (r+1) \)-iteration, \( f(S_r + \{a_k\}) = f(S_r) \) for all \( j \in A \). From 2) we know that, if Algorithm 1 do not terminate after \( r \)-iteration, there exists at least one RAP \( a_j \) that \( P_j|S_r < \tau \). Therefore, during \( (r+2) \)-iteration, choosing the RAP from \( N(a_j) \), which is not added to \( S_r \) before, will make \( f(S) \) increase. That contradicts the former conclusion. In sum, \( \Delta_{a_k} \cdot f(S_r) > 0 \).

d) Based on the analysis in the proof of Lemma 4, we can deduce the following inequation.

\[
\Delta_{a_k} \cdot f(S_r) = f(S_r + \{a_k\}) - f(S_r) \\
= \theta \cdot \sum_{v \in V} \min\{P_i|(S_r + \{a_k\}), \tau \} - \theta \cdot \sum_{v \in V} \min\{P_i|(S_r, \tau \}) \\
\geq \theta \cdot \min\{P_j|(S_r + \{a_k\}), \tau \} - \min\{P_j|(S_r, \tau \}) \\
= \theta \cdot \min\{P_j|(S_r + \{a_k\}), \tau \} - P_j|(S_r, \tau \}) \\
\geq \theta \cdot \min\{P_j|(S_r + \{a_k\}), \tau \} - P_j|(S_r, \tau \}) \\
\]

In above, \( P_j|S_r < \tau \) as Algorithm 1 doesn’t terminate after \( r \)-th iteration. If Algorithm 1 terminates after \( (r+1) \)-th iteration, \( \min\{P_j|(S_r + \{a_k\}), \tau \} = \tau \). Whereas, if Algorithm 1 doesn’t terminate after \( (r+1) \)-th iteration, \( \min\{P_j|(S_r + \{a_k\}), \tau \} = P_j|(S_r + \{a_k\}) < \tau \). In sum, \( \min\{P_j|(S_r + \{a_k\}), \tau \} \geq P_j|(S_r, \{a_k\}) \). Further, we have

\[
\Delta_{a_k} \cdot f(S_r) \geq \theta \cdot \min\{P_j|(S_r + \{a_k\}), \tau \} - P_j|(S_r) \\
\geq \theta \cdot \min\{P_j|(S_r + \{a_k\}), \tau \} - \min\{P_j|(S_r, \tau \}) \\
\geq \theta \cdot \min\{P_j|(S_r + \{a_k\}), \tau \} - P_j|(S_r, \tau \}) \\
\geq \theta \cdot \min\{P_j|(S_r + \{a_k\}), \tau \} - P_j|(S_r, \tau \}) \\
\]