This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2015.2388851, IEEE Transactions on Vehicular Technology

Prefetching-Based Data Dissemination in Vehicular Cloud Systems

Ryangsoo Kim, Hyuk Lim, and Bhaskar Krishnamachari

***Abstract*—In the last decade, vehicular ad-hoc networks (VANETs) have been widely studied as an effective method for providing wireless communication connectivity in vehicular transportation systems. In particular, vehicular cloud systems have received abundant interest for the ability to offer a variety of vehicle information services. We consider the data dissemination problem of providing reliable data delivery services from a cloud data center to vehicles through roadside wireless access points (APs) with local data storage. Due to intermittent wireless connectivity and the limited data storage size of roadside wireless APs, the question of how to use the limited resources of the wireless APs is one of the most pressing issues affecting data dissemination efficiency in vehicular cloud systems. In this paper, we devise a vehicle route-based data prefetching scheme, which maximizes data dissemination success probability in an average sense when the size of local data storage is limited and wireless connectivity is stochastically unknown. We propose a greedy algorithm and an online learning algorithm for deterministic and stochastic cases, respectively, to decide how to prefetch a set of data of interest from a data center to roadside wireless APs. Experiment results indicate that the proposed algorithms can achieve efficient data dissemination in a variety of vehicular scenarios.**

***Index Terms*—Vehicular cloud system, data dissemination, vehicular ad hoc networks, online learning, greedy algorithm, roadside wireless access point.**

I. Introduction

In the last decade, vehicular ad-hoc networks (VANETs) have been widely studied as a method for incorporating wireless communication capabilities in vehicular transportation systems for safety, energy, and comfort issues [1]. VANETs consist of two types of nodes, i.e., mobile vehicles and stationary roadside wireless access points (APs); the wireless APs serve as an infrastructure for network connectivity in VANETs. In VANETs, vehicle-to-vehicle (V2V), infrastructure-to-vehicle (I2V), and vehicle-to-infrastructure (V2I) communications are defined depending on the direction of traffic flow [2]. While V2V is primarily used for exchanging immediate driving information among neighboring vehicles on the road for safety purposes, I2V and V2I aim at data delivery services from/to the Internet through the roadside wireless APs in VANETs.

The vehicular cloud system (VCS) is a new emerging technology that can provide cloud services for various vehicle information applications such as multimedia streaming, autonomous navigation services, and remote vehicle diagnosis. The infrastructure for VCS consists of high-performance cloud servers at a data center and a number of roadside wireless APs with limited-sized local data storage. If an application requires high computational power or an extensive amount of data storage, it is more desirable to be implemented and executed as a cloud service of VCS because vehicles may have insufficient data processing and storage capability to run such a heavy application in a stand-alone manner [3]. When a vehicle needs data delivery and computing services for vehicle information applications, it uses the roadside wireless APs to contact the cloud servers. Moreover, the cloud server can perform computationally intensive tasks and disseminate output data to the vehicular subscriber. While the roadside APs are connected to the cloud servers through wired links, the connection between the vehicles and the wireless APs is intermittently available as the vehicle enters and leaves the service coverage areas of the wireless APs. Thus, owing to vehicle mobility and intermittent connectivity, data dissemination to mobile vehicles through the roadside wireless APs is a challenging problem for successful implementation of vehicular cloud systems [4].

We concentrate our attention on the problem of how to exploit the local data storages of roadside wireless APs for efficient data dissemination within the VCS. For illustration, we consider a data dissemination scenario for VCS as shown in Figure 1. Suppose a vehicle requests cloud data and the routing path is determined in advance. All of the wireless APs located on the path of the vehicle fetch the data from the cloud servers and attempt to transmit it to the vehicle when the wireless link to the vehicle is established. This approach can achieve the highest data delivery success probability, but it is not practically applicable because the size of local storage in each wireless AP is too small to store all of the data requested by multiple vehicles at the same time. Note that the number of data chunks that can be prefetched from a cloud server at a single time point is limited by the wired link capacity of wireless AP, despite the fact that the AP has a large storage capacity. For example, consider a multimedia streaming service. If a wireless AP has a gigabit Ethernet link with 128 GB solid-state drive (SSD) storage, it can prefetch 75 100-MB video clips every minute and can store 2,560 video clips. In such a case, with respect to dissemination

This research was supported in part by the U.S. National Science Foundation via award number CNS-1217260, and by the National Research Foundation of Korea grant funded by the Korea government (MSIP) (2014R1A2A2A01006002). *(Corresponding author: H. Lim.)*

R. Kim and H. Lim are with the School of Information and Communications, Gwangju Institute of Science and Technology (GIST), Gwangju 500-712, Republic of Korea. Email: {rskim,hlim}@gist.ac.kr

B. Krishnamachari is with the Department of Electrical Engineering, Viterbi School of Engineering, University of Southern California, Los Angeles, CA 90089, USA. Email: bkrishna@usc.edu

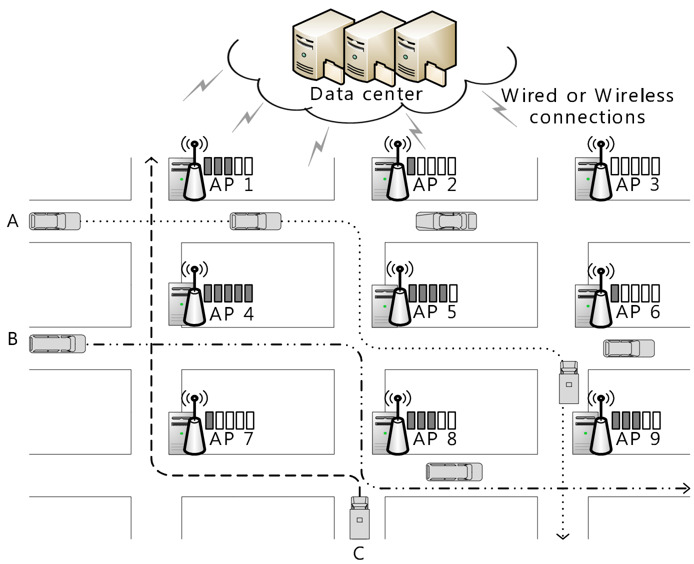


Fig. 1. Vehicle route and roadside wireless AP based data dissemination scenario for a vehicular cloud system.

efficiency, it is more desirable to fetch the data requested by more than two vehicles to the wireless APs located on the intersection of multiple paths rather than that requested by only one vehicle. For example, if the vehicles A and B request the same vehicular cloud service as shown in Figure 1, it is desirable that AP 5 has the data chunk and transmits it to the vehicles because both vehicles pass by AP 5. It is also important to consider that vehicles have connectivity to wireless APs only when they stay for at least a certain time interval within the service coverage area of the roadside APs; moreover, the connectivity is susceptible to the time-varying wireless channel.

In this paper, we devise a vehicle route-based data prefetch framework for data dissemination in vehicular cloud system, which maximizes the aggregate dissemination success probability in an average sense when the size of local data storage is limited and wireless connectivity is stochastically unknown. We formulate this data dissemination as a binary optimization problem, for which the optimal solution can be obtained by a deterministic combinatorial algorithm. We propose a greedy algorithm and analyze its approximate worst-case performance bound. We also propose an online learning algorithm based on multi-armed bandits (MABs) to maximize the aggregate dissemination success probability in an average sense by capturing the unknown stochastic characteristics of wireless connectivity. The proposed data dissemination scheme is applicable to delay-tolerant vehicular data services such as entertainment content distribution, navigation data updates, and online travel guide services.

The rest of this paper is organized as follows. First, we present a survey of related work in Section II. In Section III, we describe the system models and assumptions for data dissemination of vehicular cloud systems. In Section IV, we formulate a data dissemination problem that maximizes the aggregate dissemination success probability and propose a greedy algorithm and an online learning algorithm for deterministic and stochastic cases, respectively. In Section V, we show numerical experiment results followed by conclusions in Section VI.

II. Related Work

The data dissemination research for VANETs is summarized into the two categories of V2V and V2I/I2V communications. The data dissemination research for V2V communications focuses on how to achieve reliable and timely data delivery among mobile vehicles on roads over intermittently connected wireless links [5]–[10]. In [5], the data pouring (DP) algorithm with intersection buffering was proposed. The vehicles at intersections keep the data sent by the source node in their buffers and repeatedly rebroadcast it to other vehicles passing the intersection. In [6], the route information of vehicles, which is readily available through the GPS enabled navigation system in the vehicles, is used for alleviating channel congestion in data dissemination by selecting appropriate routing paths. In [7], the relative distance between neighboring mobile vehicles is predicted and exploited for improving reliability of data delivery. In [8], Schwartz *et al.* proposed adaptive network load control for fair data dissemination in VANETs. In [9], Ye *et al.* studied a peer-to-peer data dissemination problem and proposed a network coding based data broadcasting scheme for improving data reception efficiency. The dissemination complete time and steady-state data dissemination velocity for the peer-to-peer data dissemination were also mathematically analyzed in [9]. In [10], Sathiamoorthy *et al.* investigated vehicle-to-vehicle data sharing using erasure codes for reducing data dissemination latency in vehicular networks. In particular, they focused on the problem of how to store erasure coded data in vehicles to maximize vehicle-to-vehicle collaboration opportunities.

The data dissemination for V2I and I2V communications focuses on how to efficiently share the limited resource of roadside APs to improve the quality of data dissemination services. In [11], Liang *et al.* proposed a cooperative data dissemination approach. At the network level, network resources were cooperatively managed so as to satisfy the quality of service (QoS) requirements for realtime and non-realtime traffic. At the packet level, cooperative transmission for the sake of increasing the high packet transmission rate was proposed. In [12], rateless coding technology was applied at roadside wireless APs to improve the efficiency of data dissemination.

In I2V data dissemination, two important factors that significantly affect the data dissemination performance are the limited buffer size of roadside wireless APs and the intermittent connectivity between the wireless APs and mobile vehicles [13]–[16]. In [13], a hybrid data dissemination assisted by static nodes was proposed. When there are no vehicles that can deliver the data along a routing path, static nodes located at road intersections keep the data and forward it when the routing path becomes available. In [14], wireless transmission characteristics for sending and receiving large amounts of data from a moving vehicle to the roadside wireless APs were investigated empirically. In [15] and [16], a wireless measurement study for vehicles under different driving conditions was

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2015.2388851, IEEE Transactions on Vehicular Technology

carried out. In [17], Jeong *et al.* proposed an infrastructure-based data dissemination that utilizes the trajectory of the vehicles that the packets with a delay constraint are destined for. As a vehicle is moving along a pre-determined route path, one of the relay nodes on the path is dynamically selected as the destination for each requested data packet such that the packet delivery delay is minimized while satisfying the packet reception probability requirement. While the work in [17] mostly focused on data delivery latency rather than data dissemination efficiency for the roadside wireless APs with limited storage capacity, we consider an I2V VANET scenario where there exists a constraint on the storage size of roadside wireless APs and the network connectivity is intermittently available, and propose a deterministic greedy algorithm and an online learning-based algorithm to achieve high data dissemination efficiency in the VANETs.

The data dissemination method for VCS has some similarities with web caching strategies adopted for enhancing local access to popular Internet contents via proxy servers. Depending on user web access patterns and network topology, the web caching strategies aim to find the best place for web proxies on the network, and allow the proxies to cache popular contents for reducing user access latency and the amount of Internet access traffic [18]. In [19], Li *et al.* studied an optimal placement of web proxies among candidate sites in a tree-based network topology in order to minimize the latency for target web services. They modeled the optimal placement problem as a dynamic programming problem, and the optimal solution could be obtained in polynomial time. In [20], Qui *et al.* studied an online placement problem of web server replicas under imperfect information about client workload characteristics. They formulated the placement problem as a minimum *K*-median graph theoretic problem and devised algorithms for minimizing the client cost of accessing data replicated on web servers. In [21], Nimkar *et al.* focused on how to place multiple copies of video files on distributed proxy servers for video on demand (VoD) services, and devised greedy video placement and disk load-balancing algorithms.

The data dissemination for VCS is also related to the problem of allocating virtual computing and storage resources in large-scale distributed systems such as grid and cloud computing environments. In [22], Giurgiu *et al.* dealt with the problem of how to efficiently manage virtual network infrastructures in large scale data centers while guaranteeing resource and availability requirements. To make the optimization problem more tractable, they reduced the searching space for optimization by specifying feasible subsets of computing nodes as the candidate sites, and used a graph-based search algorithm for finding the optimal placement of virtual machines. In [23], a multi-objective ant colony system algorithm was adopted to find the optimal placement of virtual machines that can minimize the aggregate power and resource consumption in cloud infrastructures.

Our data dissemination method is similar to web caching/proxy methods in that both deal with a distributed cache/storage-based data dissemination problem. The distributed multiple cache/storage devices store a certain amount of contents to expedite the service response and to reduce the traffic amount to be downloaded from the central data server. However, there are significantly challenging issues that need to be resolved for data dissemination in vehicular network environments:

- Most web caching/proxy methods focus on maximizing the hit rate of individual proxy servers for certain content request statistics. Our VCS is formulated such that it guarantees that vehicular subscribers receive the data service successfully from at least one roadside AP while they pass by multiple APs during their traveling time.
- Web caching/proxy methods are usually designed under the assumption that proxy servers provide services to their users through a reliable link without significant loss of data. In our VCS, the vehicular subscribers communicate with roadside wireless APs through a unreliable wireless link. It is assumed that the connectivity is intermittently available, and its distribution is stochastically unknown in advance. The stochastic characteristics of wireless connectivity should be taken into consideration for reliable dissemination services.
- In these vehicular environments, the driving routes of vehicles are assumed to be predicted or obtained using online navigation. This routing information can be exploited for improving data dissemination performance. For available routing information of vehicular subscribers, our data dissemination method cooperatively manages the multiple roadside APs on the routes of the subscribers, depending on their intermittent connectivity distribution and storage capacity, in order to maximize the data service success rate for all data-vehicle pairs.

In fact, various existing web caching/proxy methods could be applied for enhancing the dissemination performance in vehicular network environments because VCS can be considered as a mobile web caching method that coordinates distributed roadside storages with network connectivity in vehicular network environment.

III. System Model

We consider a VCS that consists of cloud servers at a data center and roadside wireless APs with local data storage. Mobile vehicles have intermittent network connectivity to the cloud system through the wireless APs, which are connected to the cloud servers by means of wired infrastructure networks. To expedite data dissemination to vehicles, each AP can prefetch some data from the data center before they are requested from vehicular subscribers. We make the following assumptions for the VCS:

- The data in a cloud system is divided into a number of small chunks that are the basic units for data delivery from the data center to vehicles.
- Each AP is placed at an intersection and has limited transmission coverage such that a vehicle can download data chunks of interest only when it stays within the coverage area for at least a certain amount of time.
- Each AP has a stochastic characteristic for communicating with the mobile vehicle going through its coverage

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2015.2388851, IEEE Transactions on Vehicular Technology

TABLE I
PARAMETERS USED FOR MODELING VEHICULAR CLOUD SYSTEMS.

Parameter	Description
w	Number of wireless APs in the vehicular cloud system
v	Number of vehicles in the area of interest
u	Number of data chunks served by the vehicular cloud system
b	Maximum number of data chunks that can be stored at each local data storage
$\mathbf{R}$	Chunk request matrix $\mathbf{R} = (r_{i,j})_{v\times u}$ where $r_{i,j}$ =1 if the i-th vehicle requests the j-th chunk, and otherwise 0.
$\mathbf{S}$	Vehicle route matrix $\mathbf{S} = (s_{i,j})_{v\times w}$ where $s_{i,j}$ =1 if the i-th vehicle goes through the j-th wireless AP, and otherwise 0.
$\boldsymbol{\theta}$	Network connectivity success rate vector $\boldsymbol{\theta} = (\theta_k)_{w\times 1}$ where θ_k is an average rate of successful communications between the k-th AP and mobile vehicles.
$\mathbf{X}$	Binary decision matrix $\mathbf{X} = (x_{i,j})_{w\times u}$ where $x_{i,j}$ = 1 if the j-th chunk is to be prefetched to the i-th wireless AP and otherwise 0.
$\mathbf{P}$	Dissemination failure probability matrix $\mathbf{P} = (p_{i,j})_{v\times u}$ where $p_{i,j}$ is a probability that the i-th vehicle fails to download the j-th chunk.

region due to limited communication capacity and time-varying wireless channels.
- For effectively using a data dissemination service, the driving route of vehicles must be available in advance from online navigation and long-term archived traces.

Each vehicle that uses VCS has to report its driving route so that the data dissemination algorithm can find the best APs belonging to that route. Note that the selected APs will prefetch data chunks from a cloud server and hold the requested chunks requested by a vehicular subscriber. Whenever the route changes, it has to be reported to the VCS. If a vehicle does not have a navigation device, its driving route can be predicted using historical traces.

Given the assumptions outlined above, we first define two matrices of the chunk request matrix $\mathbf{R}$, the vehicle route matrix $\mathbf{S}$, and the network connectivity success rate vector $\boldsymbol{\theta}$ to describe a VCS. Let u, v, and w denote the numbers of chunks, vehicles, and wireless APs, respectively. $\mathbf{R} = (r_{i,j})_{v\times u}$ is a binary matrix where $r_{i,j}$ = 1 when the i-th vehicle requests the j-th chunk, and otherwise 0. $\mathbf{S} = (s_{i,k})_{v\times w}$ is also a binary matrix where $s_{i,k}$ = 1 when the i-th vehicle is expected to go through the k-th wireless AP, and otherwise 0. Last, $\boldsymbol{\theta} = (\theta_k)_{w\times 1}$ is a vector where θ_k represents the average rate of successful communication between the k-th AP and a mobile vehicle passing by the AP. We summarize the parameters used for modeling the VCS in Table I.

The data dissemination problem in this paper is to make a decision on which chunks are to be prefetched to the local data storages of wireless APs to expedite data dissemination in the VCS. Let $\mathbf{X} = (x_{i,j})_{w\times u}$ denote a binary decision matrix where $x_{i,j}$ = 1 if the j-th chunk is to be prefetched to the i-th wireless AP, and otherwise 0. For example, if $x_{i,j}$ = 1 for all i's and j's, it means that all the data chunks are distributed to every local data storage. In this case, vehicles are supposed to successfully receive as many requested chunks as possible, but this may not be efficient because all the chunks are unnecessarily copied to every local storage. In particular, the data storage capacity of roadside wireless APs is limited, and thus all of the data chunks cannot be stored in all the APs.

By using $\mathbf{R}$, $\mathbf{S}$, $\boldsymbol{\theta}$, and $\mathbf{X}$, we derive the dissemination failure probability matrix denoted by $\mathbf{P} = (p_{i,j})_{v\times u}$ where $p_{i,j}$ is the probability that the i-th vehicle fails to download the j-th chunk. Then, $p_{i,j}$ is given by the product of the probabilities that the i-th vehicle fails to download the j-th chunk at every AP that it goes through as follows:

$$p_{i,j} = r_{i,j} \times \prod_{k=1}^{w} h_{i,j,k}(s_{i,k}, \theta_k, x_{k,j}), \tag{1}$$

where $h_{i,j,k}$ is the probability that the i-th vehicle fails to download the j-th chunk at the k-th AP, and $h_{i,j,k}(s_{i,k}, \theta_k, x_{k,j}) = 1 - s_{i,k} \cdot \theta_k \cdot x_{k,j}$.

IV. PROPOSED DATA PREFETCHING SCHEME

In vehicular cloud systems, local data storages in roadside wireless APs are essential resources for expediting data dissemination from a data center to mobile vehicles. If more wireless APs that are located on the routes of a vehicle have data chunks, it is more likely that the vehicle can successfully download the data chunks. On the other hand, if too excessive data chunks are transferred to wireless APs, it is a resource waste of local data storage and may cause an increase in network delay from a data center to wireless APs. In this paper, we consider an optimization strategy that maximizes data dissemination success probability when the data storage size of each local data storage is limited. The data dissemination problem is formulated as a binary maximization over $\mathbf{X} \in \{0,1\}^{w\times u}$.

The objective of the data dissemination problem described in this paper is to maximize the dissemination success probability in the VCS. Due to the stochastic characteristics of $\boldsymbol{\theta}$ in $\mathbf{P}$, it is not possible to perfectly guarantee data chunk delivery. Instead, we maximize the dissemination success probability under the assumption that there exists a maximum boundary (or quantity) of data storage in roadside wireless APs. Let b_i denote the maximum number of data chunks that can be stored at the local data storage of the i-th AP. Note that b_i is also bounded by the maximum number of data chunks that can be downloaded from a cloud server during each decision round. Under the assumption that all APs have the same storage capacity, b_i is set to b for all $i = \{1, \cdots, w\}$. Then, we impose a constraint on the selection of the binary decision matrix such that the feasible candidates are from a finite set $\mathcal{F}_{org}$, which

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2015.2388851, IEEE Transactions on Vehicular Technology

is represented as follows:

$$\mathcal{F}_{org} = \{\mathbf{X} : \mathbf{X} \in \{0,1\}^{w \times u}, ||\mathbf{X} \cdot \mathbb{1}_u||_{max} \leq b\}, \quad (2)$$

where $\mathbb{1}_l \in \mathbb{R}^l$ is an all-ones column vector. Let $g(\mathbf{X})$ denote the aggregate dissemination success probability for all the vehicles and data chunks. Then, $g(\mathbf{X})$ is given by

$$g(\mathbf{X}) = \sum_{i=1}^{v} \sum_{j=1}^{u} r_{i,j} \cdot (1 - p_{i,j}). \quad (3)$$

We also define a set function $g_a(\mathcal{A}_\mathbf{X})$ for $g(\mathbf{X})$, i.e., $g_a(\mathcal{A}_\mathbf{X}) = g(\mathbf{X})$ where $\mathcal{A}_\mathbf{X} = \{(i,j) : x_{i,j} = 1, 1 \leq i \leq w, 1 \leq j \leq u\}$ is the index set for $x_{i,j} = 1$ in $\mathbf{X}$. The optimal binary decision matrix $\mathbf{X}^*$ that maximizes $g(\mathbf{X})$ is to be selected from $\mathcal{F}_{org}$ as follows:

$$\mathbf{X}^* = \arg \max_{\mathbf{X} \in \mathcal{F}_{org}} g(\mathbf{X}). \quad (4)$$

Algorithm 1 Deterministic greedy data dissemination for VCS

1: // Initialization
2: $\mathbf{X}_g^* = \arg\ \max_{\mathbf{X} \in \{\mathbf{X}:\mathbf{X} \in \{0,1\}^{w \times u}, |\mathcal{A}_\mathbf{X}|=3\}} g(\mathbf{X})$
3: $\mathcal{A}_{\mathbf{X}^C} = \phi$;
4: // Main loop
5: **for** $l = 4$ to $w \times b$ **do**
6: $\quad \{(y,z)\} = \arg\ \max_{\{(i,j)\} \in \mathcal{A}_{\mathbb{1}_{w \times u}} \setminus (\mathcal{A}_{\mathbf{X}_g^*} \cup \mathcal{A}_{\mathbf{X}^C})}$
7: $\quad \left(g_a(\mathcal{A}_{\mathbf{X}_g^*} \cup \mathcal{A}_{x_{i,j}}) - g_a(\mathcal{A}_{\mathbf{X}_g^*}) \right)$;
8: $\quad$ **if** $(\mathbf{X}_g^* \mathbb{1}_u)_y >= b$ **then**
9: $\quad\quad \mathcal{A}_{\mathbf{X}^C} = \mathcal{A}_{\mathbf{X}^C} \cup \mathcal{A}_{x_{y,z}}$;
10: $\quad\quad$ Go back to line 6.
11: $\quad$ **end if**
12: $\quad \mathcal{A}_{\mathbf{X}_g^*} = \mathcal{A}_{\mathbf{X}_g^*} \cup \mathcal{A}_{x_{y,z}}$;
13: $\quad \mathbf{X}_g^*(y,z) = 1$;
14: **end for**

A. Deterministic greedy data dissemination for VCS

For simplicity, we first assume that the stochastic characteristics of the network connectivity success rate $\boldsymbol{\theta}$ between APs and vehicles are completely observed, thereby being able to exploit the deterministic statistics for data dissemination. In other words, the values $\boldsymbol{\theta} = [\theta_1, \cdots, \theta_w]$ are completely known to the data center in advance. Then, the data dissemination problem is a binary optimization problem, and its optimal solution can be obtained by a deterministic combinatorial algorithm. One may attempt to solve this binary optimization by a brute-search algorithm, which enumerates all possible candidates and checks whether each candidate satisfies the problem statement. However, the complexity grows exponentially with respect to the dimension of $\mathbf{X}$, and thus it is not scalable and limited in applicability to small vehicular systems with only a few tens of $(w \times u)$. To make this binary optimization problem more tractable, we derive the following proposition and then propose a greedy algorithm based on the proposition.

Proposition 1: For maximizing data dissemination success probability, a finite set of binary decision matrices can be reduced as follows:

$$\mathcal{F} = \{\mathbf{X} : \mathbf{X} \in \{0,1\}^{w \times u}, \mathbf{X} \cdot \mathbb{1}_u = b \cdot \mathbb{1}_w\}.$$

Proof: Suppose that an element $x_{y,z} \in \mathbf{X}$ is 0. When $x_{y,z}$ is changed from 0 to 1, the variation of the cost function $g(\mathbf{X})$ in (3) with respect to $x_{y,z}$ is $\sum_{i=1}^{v} r_{i,z} \times \prod_{k=1}^{w \setminus y}(1 - s_{i,k} \cdot \theta_k \cdot x_{k,z}) \times (s_{i,y} \cdot \theta_y)$, which is greater than or equal to 0. This implies that as more elements in $\mathbf{X}$ change their value to 1, the variation of $g(\mathbf{X})$ is greater than or equal to 0 (the equality holds when $r_{i,z}$ or $s_{i,y}$ is equal to 0 for all $i \in \{1, \cdots, v\}$). This means that the cost function $g(\mathbf{X})$ increases over $\mathbf{X}$. Therefore, the cost function in (4) can be maximized with the largest number of 1's in $\mathbf{X}$ that satisfies $\mathbf{X} \cdot \mathbb{1}_u = b \cdot \mathbb{1}_w$ in the constraint. ■

Based on the above proposition, we propose a greedy algorithm that iteratively finds the sub-optimal solution on $\mathcal{F}$ by setting one element of $\mathbf{X}$ to 1 at each iteration. The detailed procedure is given in Algorithm 1. The algorithm starts with $\mathbf{X}_g^*$, which is one that gives the largest dissemination success probability among all the feasible $\mathbf{X}$'s with $|\mathcal{A}_\mathbf{X}| = 3$ on line 2. [1] Then, at each iteration, the algorithm picks one element $x_{y,z} \in \mathcal{A}_{\mathbb{1}_{w \times u}} \setminus (\mathcal{A}_{\mathbf{X}_g^*} \cup \mathcal{A}_{\mathbf{X}^C})$ maximizing the increment of dissemination success probability, i.e., $g_a(\mathcal{A}_{\mathbf{X}_g^*} \cup \mathcal{A}_{x_{y,z}}) - g_a(\mathcal{A}_{\mathbf{X}_g^*})$ where $\mathbb{1}_{w \times u}$ is an all-ones $w \times u$ matrix, and $\mathcal{A}_{\mathbf{X}^C}$ is the index set for all $x_{i,j}$s in $\mathbf{X}_g^*$ that are 0 but are prohibited from being converted to 1 due to the limited storage capacity of the roadside wireless APs. On line 7, if the number of ones in the y-th row of $\mathbf{X}_g^*$ is less than b (i.e., $(\mathbf{X}_g^* \mathbb{1}_u)_y < b$ where $(\mathbf{a})_y$ represents the y-th element of vector $\mathbf{a}$), then we set the $x_{y,z}$ in $\mathbf{X}_g^*$ to 1 (i.e., $\mathcal{A}_{\mathbf{X}_g^*} = \mathcal{A}_{\mathbf{X}_g^*} \cup \mathcal{A}_{x_{y,z}}$). Otherwise, we set $x_{y,z}$ in $\mathbf{X}_g^*$ to 0 (i.e., $\mathcal{A}_{\mathbf{X}^C} = \mathcal{A}_{\mathbf{X}^C} \cup \mathcal{A}_{x_{y,z}}$) and repeat the above steps. This algorithm terminates when the number of ones in each row of $\mathbf{X}_g^*$ is equal to b.

In Algorithm 1, the number of iterations is equal to $(w \times b - 3)$ because each local data storage in the wireless AP has b number of 1's in its corresponding $\mathbf{X}$'s row. In each iteration inside the loop, the maximum from the vector of $(w \times u)$ needs to be searched. The evaluation of $g_a(\cdot)$ incurs the complexity of $O(v \times u)$ from (3). As a result, the computational complexity of Algorithm 1 becomes $O(w^2 u^2 bv)$, which corresponds to quadratic complexity with respect to w and u. Even though we assume that all the wireless APs have the same data storage capacity b in (2), Algorithm 1 can be easily extended to the case in which the capacities are not the same without any significant modification. Because the storage capacity of the i-th AP is b_i, the number of iterations for finding the optimal solution is set to $(\sum_{i=1}^{w} b_i - 3)$. Note that the local data storage of the i-th AP has b_i number of 1's in its corresponding $\mathbf{X}$'s row.

As greedy algorithms may fail to find the globally optimal solution, it is necessary to check the worst-case performance of a greedy algorithm by checking its approximation factor. If its approximation is bounded by a constant factor, the greedy algorithm is capable of finding the sub-optimal solution in polynomial-time. To derive its approximation, we derive the following proposition.

[1]Note that the reason for enumerating all the feasible $\mathbf{X}$'s with $|\mathcal{A}_\mathbf{X}| = 3$ in the first iteration is to make it possible to derive an approximation of the proposed greedy algorithm, as described in Appendix A.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2015.2388851, IEEE Transactions on Vehicular Technology

Proposition 2: The proposed dissemination problem in (4) can be transformed into a submodular maximization problem (SMP).

Proof: Submodularity is an intuitive notion of *diminishing returns*, which implies that adding an element to a small set gives more returns than adding that same element to a larger set [24]. It is defined as follows: A real-valued set function H, defined on subsets of a finite set $\mathcal{S}$, is called *submodular* if it satisfies

$$H(\mathcal{B}_1 \cup s) - H(\mathcal{B}_1) \geq H(\mathcal{B}_2 \cup s) - H(\mathcal{B}_2) \quad (5)$$

for all $\mathcal{B}_1 \subseteq \mathcal{B}_2 \subseteq \mathcal{S}$ and for all $s \in \mathcal{S} \setminus \mathcal{B}_2$. To verify submodularity, we consider the problem in (4) as follows:

$$\max_{\mathbf{X} \in \mathcal{F}} g_a(\mathcal{A}_{\mathbf{X}}) \text{ subject to } \mathbf{X}\mathbb{1}_u = b \cdot \mathbb{1}_w. \quad (6)$$

Note that $g_a(\mathcal{A}_{\mathbf{X}})$ is also a nondecreasing set function because $g(\mathbf{X})$ is a nondecreasing function over $\mathbf{X}$ as shown in Proposition 1. Consider $\mathcal{A}_{\mathbf{X}_1} \subseteq \mathcal{A}_{\mathbf{X}_2} \subseteq \mathcal{A}_{\mathbb{1}_{v\times u}}$ and $\mathcal{A}_{x_{y,z}} \in \mathcal{A}_{\mathbb{1}_{w\times u}} \setminus \mathcal{A}_{\mathbf{X}_2}$. Then, we have the following:

$$\begin{aligned} & g_a(\mathcal{A}_{\mathbf{X}_1} \cup \mathcal{A}_{x_{y,z}}) - g_a(\mathcal{A}_{\mathbf{X}_1}) \\ &= \sum_{i=1}^{v} r_{i,z} \cdot s_{i,y} \cdot \theta_y \cdot x_{y,z} \cdot \prod_{k \in \{k | (k,z) \in \mathcal{A}_{\mathbf{X}_1}\}} (1 - s_{i,k} \cdot \theta_k \cdot x_{k,z}), \end{aligned} \quad (7)$$

where the right-hand side is greater than or equal to zero. Then, we can show the following inequalities:

$$\begin{aligned} & \{g_a(\mathcal{A}_{\mathbf{X}_1} \cup \mathcal{A}_{x_{k',j'}}) - g_a(\mathcal{A}_{\mathbf{X}_1})\} - \{g_a(\mathcal{A}_{\mathbf{X}_2} \cup \mathcal{A}_{x_{k',j'}}) - g_a(\mathcal{A}_{\mathbf{X}_2})\} \\ &= \sum_{i=1}^{v} r_{i,j'} \cdot s_{i,k'} \cdot \theta_{k'} \cdot x_{k',j'} \cdot \prod_{k \in \{k | (k,j') \in \mathcal{A}_{\mathbf{X}_1}\}} (1 - s_{i,k} \cdot \theta_k \cdot x_{k,j'}) \\ & \quad \cdot (1 - \prod_{k \in \{k | (k,j') \in \mathcal{A}_{\mathbf{X}_2} \setminus \mathcal{A}_{\mathbf{X}_1}\}} (1 - s_{i,k} \cdot \theta_k \cdot x_{k,j'})) \geq 0. \end{aligned} \quad (8)$$

According to (8), the function $g_a(\mathcal{A}_{\mathbf{X}})$ is a sub-modular set function, and the problem in (6) is an SMP. ■

Based on the above proposition, the greedy algorithm is able to achieve a constant factor $(1 - e^{-1})$ approximation of the optimal value of (4). The detailed procedure to achieve such an approximation bound is described in Appendix A. However, in practice, it is desirable to deal with the network connectivity success rate $\boldsymbol{\theta}$ as a random process because its statistical property is unknown in advance. In such a case, it is not possible to directly determine how many data chunks are prefetched to the local data storages. In the next subsection, we propose an alternative way to observe and exploit the stochastic characteristics of $\boldsymbol{\theta}$ for maximizing the dissemination success probability in an average sense.

B. Online learning-based data dissemination algorithm

In this paper, we adopt the stochastic multi-armed bandit (MAB) based online learning framework presented in [25] to solve the data dissemination problem in (4). MABs are widely used to solve combinatorial optimization problems for cost functions with unknown random variables. The MAB framework gradually learns the stochastic characteristics of random variables with unknown distribution and then determines an optimal policy to maximize the cost function in an average sense. The performance of the MAB is evaluated by analyzing the *regret*, which is defined as the aggregated difference between the maximum costs given by a globally optimal solution and those by the MAB over time. If the *regret* increases sub-linearly, it implies that the solution of the MAB gradually converges to a globally optimal solution in a certain number of iterations. In this subsection, we propose an MAB-based online learning algorithm for the data dissemination and perform the *regret* analysis to show that the solution of our proposed algorithm converges to a globally optimal solution.

1) Policy design: In our data dissemination problem, the network connectivity success rate $\boldsymbol{\theta}$ for roadside wireless APs is a random variable with unknown distribution that changes over time. Let n be a time index representing a decision period for online learning iterations and $\mathbf{t} = (t_k)_{w\times 1}$ denote the random variables representing network connectivity of the APs, where $\theta_k = \mathbb{E}[t_k]$ for all $k = \{1, \cdots, w\}$. The proposed MAB-based online learning algorithm measures the mean network connectivity success rate $\boldsymbol{\theta}$ at each decision period and finds an optimal solution on $\mathcal{F}$ that maximizes a cost function with mean network connectivity success rate. The globally optimal binary decision matrix $\mathbf{X}^*$ is given by

$$\mathbf{X}^* = \arg\max_{\mathbf{X} \in \mathcal{F}} \sum_{i=1}^{v} \sum_{j=1}^{u} r_{i,j} \cdot (1 - \prod_{k=1}^{w} h_{i,j,k}(s_{i,k}, \theta_k, x_{k,j})). \quad (9)$$

The detailed procedure is given in Algorithm 2. The idea for this algorithm was inspired by algorithm CWF2 in [25], which exploits the information gained from the operation of each action to determine a dependent action and achieves a logarithmically growing regret.

In Algorithm 2, the initial learning process is performed for each AP, so that at least one data chunk may be downloaded from the AP to vehicles. On lines 3 – 8, at the p-th iteration, an arbitrary binary decision matrix $\mathbf{X} \in \mathcal{F}$ is chosen such that the number of data chunks downloaded from the p-th AP, which is $(\{(\mathbf{S}^T\mathbf{R}) \odot \mathbf{X}\}\mathbb{1}_u)_p$, is greater than or equal to 1 in order to measure and estimate the initial values of the instantaneous network connectivity success rate θ'_k and the accumulated mean network connectivity success rate $\overline{\theta}_k$. Subsequently, the selected arm $\mathbf{X}(n)$ is played, and θ'_k and $\overline{\theta}_k$ are measured and updated. The instantaneous network connectivity success rate $\boldsymbol{\theta}' = [\theta'_1, \cdots, \theta'_w]^T$ is given by

$$\theta'_k = \frac{\sum_{i=1}^{v} \sum_{j=1}^{u} r_{i,j} \cdot (1 - h_{i,j,k}(s_{i,k}, t_k, x_{k,j}(n)))}{(\{(\mathbf{S}^T\mathbf{R}) \odot \mathbf{X}(n)\}\mathbb{1}_u)_k}. \quad (11)$$

The accumulated mean network connectivity success rate $\overline{\boldsymbol{\theta}} = [\overline{\theta}_1, \cdots, \overline{\theta}_w]^T$ is updated as follows:

$$\overline{\theta}_k = \frac{\overline{\theta}_k \cdot m_k + \theta'_k \cdot (\{(\mathbf{S}^T\mathbf{R}) \odot \mathbf{X}(n)\}\mathbb{1}_u)_k}{m_k + (\{(\mathbf{S}^T\mathbf{R}) \odot \mathbf{X}(n)\}\mathbb{1}_u)_k}, \; k = \{1, \cdots, w\}, \quad (12)$$

where $\mathbf{m} = [m_1, \cdots, m_w]^T$ is the number of observation times up to the current iteration for the APs. Based on $\overline{\theta}$, an optimal binary decision matrix is determined as described in (10) on line 12. The proposed online learning algorithm

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2015.2388851, IEEE Transactions on Vehicular Technology

Algorithm 2 Proposed online learning algorithm

// Initialization
$n = 0$;
for $p = 1$ to w **do**
  $n := n + 1$;
  Play any arm $\mathbf{X} \in \mathcal{F}$ such that $(\{(\mathbf{S}^T\mathbf{R}) \odot \mathbf{X}\}\mathbb{1}_u)_p \geq 1$;
  $\theta'_k = \frac{\sum_{i=1}^{v}\sum_{j=1}^{u} r_{i,j}\cdot(1-h_{i,j,k}(s_{i,k},t_k,x_{k,j}(n)))}{(\{(\mathbf{S}^T\mathbf{R})\odot\mathbf{X}(n)\}\mathbb{1}_u)_k}, \quad k = \{1, \cdots, w\}$
  $\overline{\theta}_k = \frac{\overline{\theta}_k\cdot m_k + \theta'_k\cdot(\{(\mathbf{S}^T\mathbf{R})\odot\mathbf{X}(n)\}\mathbb{1}_u)_k}{m_k + (\{(\mathbf{S}^T\mathbf{R})\odot\mathbf{X}(n)\}\mathbb{1}_u)_k}, \quad m_k = m_k + (\{(\mathbf{S}^T\mathbf{R}) \odot \mathbf{X}(n)\}\mathbb{1}_u)_k, \quad k = \{1, \cdots, w\}$;
end for
// Main loop
while 1 **do**
  $n := n + 1$;
  Play any arm $\mathbf{X} \in \mathcal{F}$ which solves the following maximization problem:

$$\max_{\mathbf{X}\in\mathcal{F}} \sum_{i=1}^{v}\sum_{j=1}^{u} r_{i,j} \cdot \left(2 - \prod_{k=1}^{w} h_{i,j,k}(s_{i,k}, \overline{\theta}_k, x_{k,j}) - \prod_{k=1}^{w} \max\{h_{i,j,k}(s_{i,k}, \sqrt{\frac{(w+1)\ln n}{m_k}}, x_{k,j}), 0\}\right); \tag{10}$$

  $\theta'_k = \frac{\sum_{i=1}^{v}\sum_{j=1}^{u} r_{i,j}\cdot(1-h_{i,j,k}(s_{i,k},t_k,x_{k,j}(n)))}{(\{(\mathbf{S}^T\mathbf{R})\odot\mathbf{X}(n)\}\mathbb{1}_u)_k}, \quad k = \{1, \cdots, w\}$
  $\overline{\theta}_k = \frac{\overline{\theta}_k\cdot m_k + \theta'_k\cdot(\{(\mathbf{S}^T\mathbf{R})\odot\mathbf{X}(n)\}\mathbb{1}_u)_k}{m_k + (\{(\mathbf{S}^T\mathbf{R})\odot\mathbf{X}(n)\}\mathbb{1}_u)_k}, \quad m_k = m_k + (\{(\mathbf{S}^T\mathbf{R}) \odot \mathbf{X}(n)\}\mathbb{1}_u)_k, \quad k = \{1, \cdots, w\}$;
end while

iteratively finds a globally optimal binary decision matrix that maximizes the aggregate dissemination success probability in an average sense. Note that $\overline{\theta}_k$ gradually converges to the *actual* mean network connectivity success rate as $\overline{\theta}_k$ is updated over time. The proposed online learning algorithm needs two storage units of size $w \times 1$ to store $\overline{\boldsymbol{\theta}}$ and the number of observation times $\mathbf{m}$.

2) Regret analysis: We perform the *regret* analysis to show that the solution of the proposed online learning algorithm converges to a globally optimal solution in a certain number of iterations. The *regret* of the proposed algorithm is an aggregate discrepancy between the maximum aggregate dissemination success probabilities by a globally optimal solution and by the proposed algorithm in Algorithm 2. The *regret* after N iterations is given by

$$\mathcal{R}(N) = N \cdot g(\mathbf{X}^*) - \sum_{n=1}^{N} g(\mathbf{X}(n)), \tag{13}$$

where $g(\mathbf{X}^*) = \max_{\mathbf{X}\in\mathcal{F}} g(\mathbf{X})$ is the maximum aggregate dissemination success probability by the optimal binary decision matrix $\mathbf{X}^*$.

The *regret* analysis derives the upper bound of the *regret* after N iterations. The upper bound can be obtained as a function of the upper bound of the number of times for which a non-optimal binary decision matrix is selected. Let $T_{NO}(N)$ denote the number of times for which a non-optimal binary decision matrix is selected for the first N iterations. To show the upper bound of $T_{NO}(N)$, we define $T_k(N)$ as a counter for the k-th wireless AP. Once the online learning algorithm selects a non-optimal binary decision matrix, the index j such that $j = \arg\min_{k\in\{1,\cdot,w\}} m_k$ is selected, and the corresponding counter $T_j(N)$ is increased by 1. If there are more than two indexes, one index is selected arbitrarily and the corresponding counter is increased by 1. Then, only one counter will increment its value when the non-optimal binary decision matrix is selected, and the following equation must hold:

$$T_{NO}(N) = \sum_{k=1}^{w} T_k(N).$$

Based on the above equation, the upper bound of *regret* is given by

$$\mathcal{R}(N) \leq \Delta_{\max} \cdot \left(\sum_{k=1}^{w} T_k(N)\right), \tag{14}$$

where $\Delta_{\max} = g(\mathbf{X}^*) - \min_{\mathbf{X}\in\mathcal{F}} g(\mathbf{X})$. Under the above inequality, an upper bound of the *regret* function $\mathcal{R}(N)$ is determined by the upper bound of the counter $T_k(N)$, which is given as follows:

$$\mathbb{E}[T_k(N)] \leq \frac{(w+1)\ln N}{\theta_{\min}^2} + 1 + \frac{\pi^2}{3}(w^2 b), \tag{15}$$

where $\theta_{\min}$ is a constant less than or equal to 1. The detailed derivation of the upper bound for the counter $T_k(N)$ is described in Appendix B. The upper bound of the *regret* is as follows:

$$\mathcal{R}(N) \leq \Delta_{\max} \cdot \left(\frac{w(w+1)\ln N}{\theta_{\min}^2} + w + \frac{\pi^2}{3}(w^3 b)\right). \tag{16}$$

As shown in the above equation, the *regret* function $\mathcal{R}(N)$ increases sub-linearly with respect to the number of iterations. This sub-linear increase implies that the optimal solution of the proposed algorithm gradually converges to a globally optimal binary decision matrix after a certain number of iterations.

3) Comparison with other learning algorithms: While the proposed algorithm was inspired by the CWF2 algorithm in [25], there are two major enhancements in our proposed algorithm. First, the cost function of the optimization problem in (10) designed for finding the optimal data dissemination

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2015.2388851, IEEE Transactions on Vehicular Technology

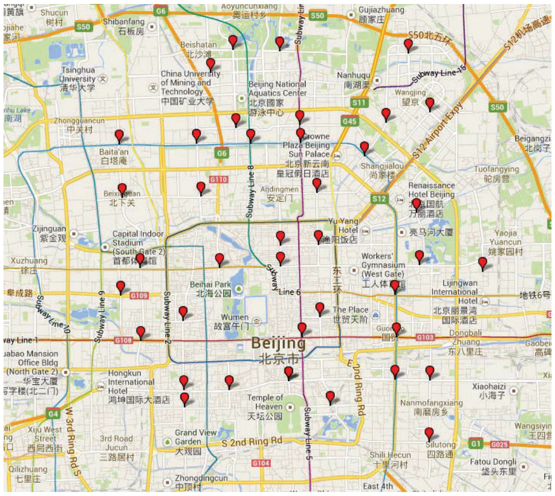


Fig. 2. Map of the roadside wireless APs randomly deployed at the intersections of roads in Beijing.

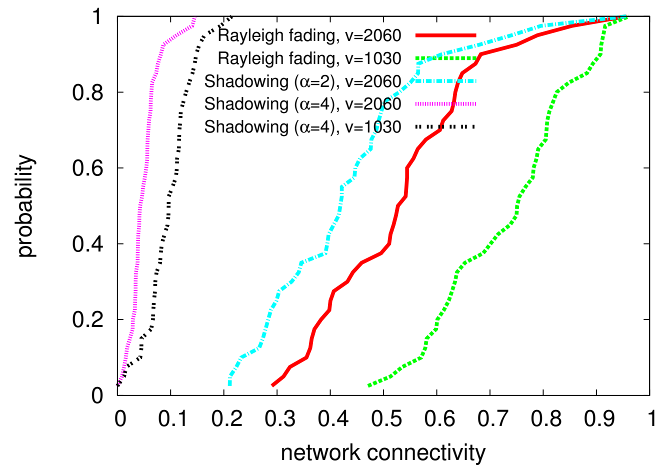


Fig. 3. Simulation results of network connectivity for different propagation models and vehicular subscriber densities.

strategy is too complicated to be solved by the CWF2 algorithm. This is because the CWF2 algorithm deals with a cost function that consists of a weighted linear combination where each term is a function of a single random variable. On the other hand, our algorithm deals with a more complicated cost function with nonlinear dependencies that include the multiplication of unknown random variables and linear reward terms. Other online learning algorithms such as UCB1 [26] and LLR [27] are limitedly applicable when the cost function is a function of one random variable and is a weighted linear combination of random variables, respectively.

Second, in the CWF2 algorithm, each arm is associated with one unknown random variable, and thus each unknown random variable can be explored only one time in each decision period when the online learning algorithm decides to exploit the corresponding arm. In contrast, our online learning algorithm can explore multiple unknown random variables several times in one decision period because each unknown random variable is associated with multiple arms. As a result, our online learning algorithm finds the features of unknown random variables more rapidly than the CWF2 algorithm.

V. Performance Evaluation

A. Experiment Environment

In this section, we present the results of real-life vehicle-trace based experiments designed to evaluate the efficiency of the proposed data dissemination methods for VCSs. It is assumed that the roadside wireless APs are deployed on the intersections of roads, and the vehicular subscribers drive along routes, which are guided by navigation system and are provided to the cloud service center when the vehicular subscribers request their cloud services. The vehicular subscribers can communicate with the APs located nearby at the intersections on their route.

For vehicle traffic trace, we use GPS traces of 2,060 taxis in Beijing as done in [28].[2] We randomly deploy 40 roadside wireless APs at the intersections as shown in Figure 2. The vehicular route matrix **S** is obtained from the route of vehicular subscribers where $s_{i,j} = 1$ if the i-th vehicular subscriber goes through the intersection at which the j-th AP is deployed. For data dissemination, it is assumed that the data chunks have the same size and the data rates of the APs are identical. The APs access the wireless channel using IEEE 802.11 DCF. The network connectivity success rate $\boldsymbol{\theta} = [\theta_1, \cdots, \theta_w]$ has a positive value in the range of [0, 1]. The chunk request matrix **R** is randomly set to either 0 or 1 under the assumption that the vehicular subscribers randomly request data chunks.

We used the *ns*-2 network simulator to characterize the network connectivity distribution of the wireless APs. The wireless channels of the links from wireless APs to vehicular subscribers are modeled as either a Rayleigh fading or shadowing model with a path-loss exponent. The path-loss exponent for the wireless channel varies from 2 to 4. Figure 3 shows the simulation results for network connectivity of the 40 wireless APs for different wireless link models and vehicular subscriber density. The cumulative density function (CDF) in Figure 3 indicates that the network connectivity depends heavily on both wireless link characteristics and the number of vehicles passing by the APs. For each wireless link model, the network connectivity becomes worse when the number of vehicles increases. In the shadowing model, the network connectivity becomes worse when the path-loss exponent increases because the communication range of APs decreases due to the significant signal attenuation.

B. Experiment Results for Deterministic Greedy Data Dissemination for VCS

In this subsection, we focus on the deterministic greedy data dissemination algorithm and numerically evaluate its performance in various data dissemination scenarios. Each

[2]This vehicle-trace dataset was obtained from research conducted by the University of Southern California's Autonomous Networks Research Group, http://anrg.usc.edu.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2015.2388851, IEEE Transactions on Vehicular Technology

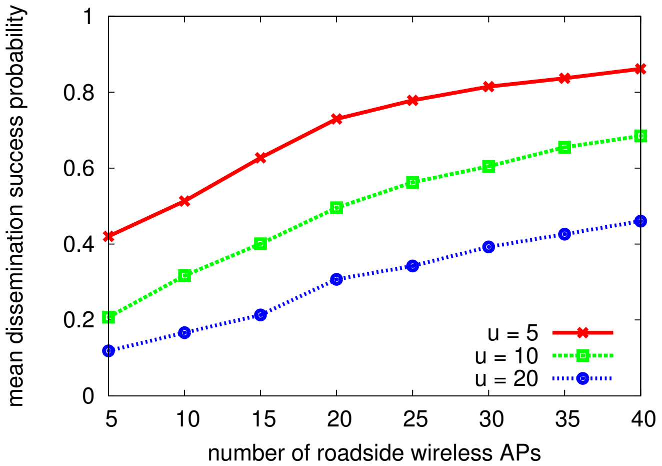


Fig. 4. Mean dissemination success probability of all vehicle-data chunk pairs with respect to the number of roadside wireless APs (Rayleigh fading, v = 2060).

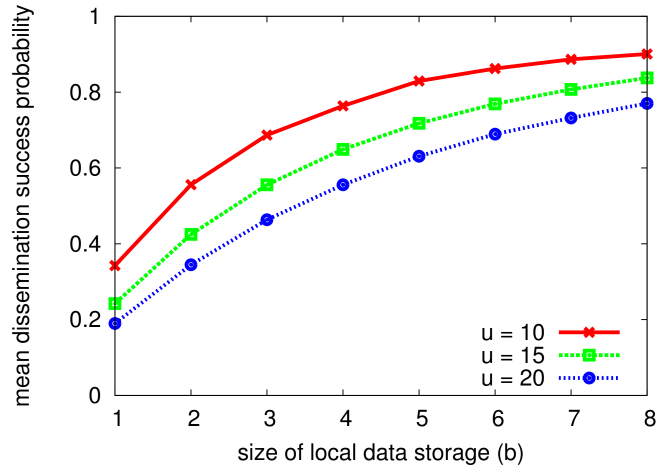


Fig. 6. Mean dissemination success probability of all vehicle-data chunk pairs with respect to the size of local data storage (Rayleigh fading, v = 2060).

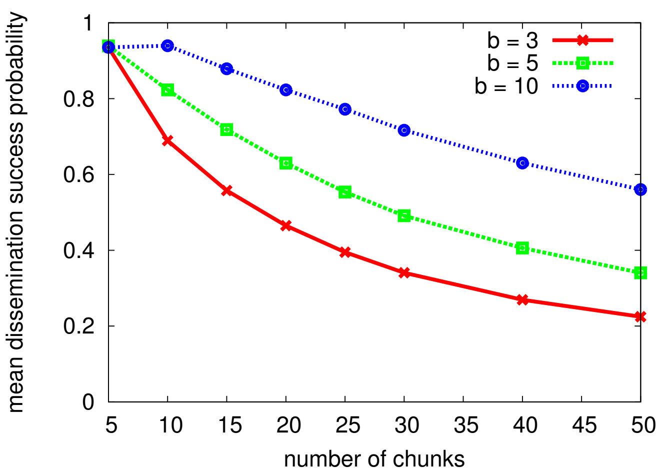


Fig. 5. Mean dissemination success probability of all vehicle-data chunk pairs with respect to the number of data chunks (Rayleigh fading, v = 2060).

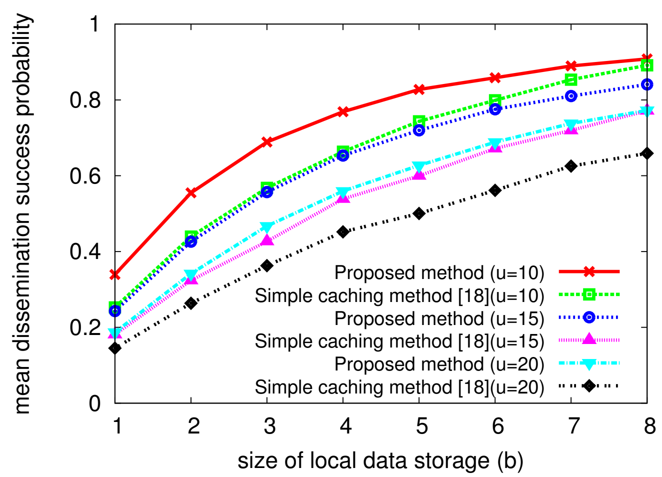


Fig. 7. Mean dissemination success probability of all vehicle-data chunk pairs with respect to the storage size for the proposed and the simple caching methods (Rayleigh fading, v = 2060).

point in Figures 4-9 is the average value of 10 runs for different $\mathbf{R}$'s.

Figure 4 depicts the mean dissemination success probability for different numbers of data chunks with respect to the number of roadside wireless APs, for the case when the roadside wireless APs are randomly chosen from the 40 APs depicted in Figure 2 and the size of the local data storage is set to 3 (b = 3). In Figure 4, we observe that as more APs are deployed on the intersections of the road, the performance of the data dissemination algorithm is enhanced. This is because the vehicular subscribers have more opportunities to download the requested data chunks due to the increasing number of APs located on the route of the vehicular subscribers. This result implies that the proposed algorithm will continue to achieve better data dissemination performance by employing more roadside wireless APs for the VCS.

Figure 5 depicts the average value of dissemination success probability for different sizes of data storage with respect to the number of data chunks served by the vehicular cloud servers. It shows that the data dissemination performance is degraded when the number of data chunks served by the cloud servers increases. The reason for this deterioration is that the vehicles request more chunks that are diverse while the capacity of local storage at each AP is limited. In order to improve data dissemination performance, the size of the local data storage needs to be increased, and the number of data chunks served by the VCS needs to be decreased.

Figure 6 depicts the average probability of dissemination success with respect to the size of local data storage. In Figure 6, the mean probability of dissemination success increases as the size of local data storage increases. This is because the roadside wireless APs can store more data chunks requested by the vehicular subscribers as the size of local data storage increases. However, the increase of data storage size incurs additional costs for the VCS infrastructure. It implies that there is trade-off between the infrastructure expenditure for data storage capacity and the data dissemination performance in VCS.

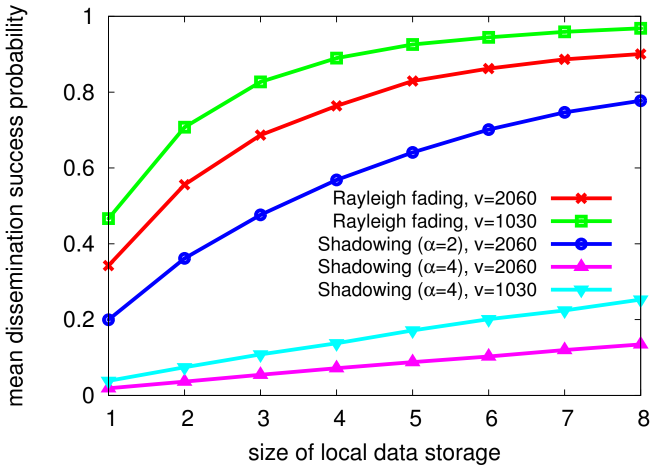


Fig. 8. Mean dissemination success probability of all vehicle-data chunk pairs with respect to the storage size for different wireless channel models and vehicular subscriber densities (u = 20).

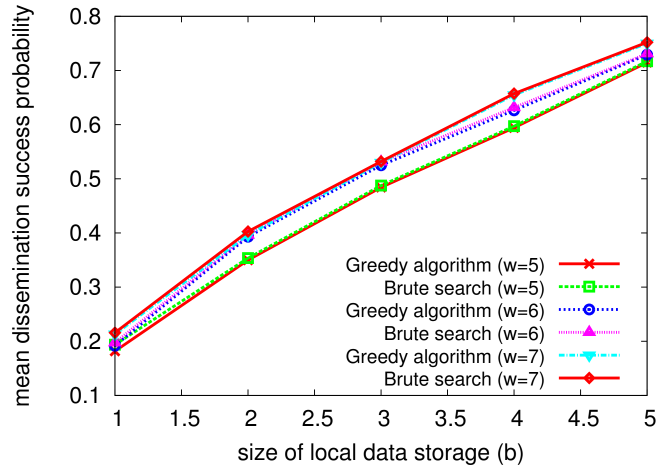


Fig. 9. Mean dissemination success probability of all vehicle-data chunk pairs with respect to the data storage size for the proposed greedy and the brute search algorithms (Rayleigh fading, v = 2060, u = 6).

For comparing the proposed data dissemination solution with web caching/proxying methods, we implemented a simple caching method. In general, most caching/proxy algorithms have a caching policy to make each proxy server keep the most popular contents so as to maximize the average hit rate of the contents requested by the users. The simple replica placement method was implemented to allow each AP to identify and keep the b chunks with the largest number of requests at each AP. In Figure 7, as the capacity of data storage for roadside wireless APs increases, both dissemination schemes achieve better dissemination performance. In the entire range, the proposed dissemination scheme outperforms the simple caching method. Under the simple caching method, each AP knows the accurate chunk request information of vehicular subscribers, but it does not consider whether its neighboring APs have the same copy of the chunks in their buffers. Furthermore, the simple caching method does not reflect the different network connectivity statistics for the caching decision. This experiment result indicates that it is important to exploit the information about the prefetching of data chunks in neighboring APs and the non-identical wireless connectivity statistics in the data dissemination decision.

We next evaluated data dissemination performance with respect to the storage size for the wireless channel models and vehicular subscriber densities. In Figure 8, the dissemination success probability is higher when the number of vehicular subscribers v is smaller for both two channel models. In addition, for the shadowing model, the dissemination performance is better when the path-loss exponent α is smaller. The reason is that, as shown in Figure 3, the network connectivity increases when v and α decrease. This implies that if the network connectivity between vehicular subscribers and roadside wireless APs increases, the dissemination method achieves better dissemination performance.

Finally, we verified the sub-optimality of the proposed greedy data prefetching algorithm. We compared the dissemination success probabilities of the greedy based prefetching algorithm with those of a brute search algorithm. Note that the brute search algorithm enumerates all the possible data prefetching cases and then chooses the optimal data prefetching decision that gives maximum data dissemination performance. Figure 9 depicts the average data dissemination success probabilities of the proposed greedy algorithm and brute search algorithm with respect to the data storage size, for the case in which the number of data chunks served by the cloud servers is 6. Figure 9 shows that the dissemination performance of the proposed greedy algorithm is almost the same as that of the brute search algorithm. This result implies that the proposed greedy algorithm near-optimally distributes the data chunks to the local storage of the APs with reasonable computational complexity as compared to the brute search algorithm.

C. Experiment Results for Online Learning-Based Data Dissemination

In this subsection, we focus on the proposed online learning-based data dissemination algorithm and numerically evaluate its performance in a VCS. For this experiment, we set the number of the data chunks served by the cloud servers to 5, and the data storage size is set to 3 (u = 5, b = 3). The proposed online learning algorithm iteratively estimates the network connectivity success rate $\boldsymbol{\theta}$.

For the performance evaluation, we compare the regret value of the proposed algorithm with that of a naive online learning approach using the UCB1 policy in [26]. The naive online learning approach is designed to learn the aggregate data dissemination success probability for all possible data prefetching sets. Note that the proposed algorithm learns the network connectivity characteristics themselves rather than the aggregated data dissemination probabilities. Table II represents the data storage size required to store the information learned and updated by the algorithms. Under the UCB1 policy, the data prefetching set that maximizes $\overline{Y}_k + \sqrt{2 \ln n / m_k}$ is selected at each decision period for prefetching data chunks to the APs, where n is the number of decision periods, $\overline{Y}_k$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2015.2388851, IEEE Transactions on Vehicular Technology

TABLE II
SIZE OF DATA STORAGES TO STORE THE INFORMATION LEARNED BY THE PROPOSED ALGORITHM AND THE NAIVE APPROACH UCB1.

Parameters	Proposed algorithm	Naive approach
$w = 5, u = 5, b = 3$	5	1.024×10^8
$w = 6, u = 5, b = 3$	6	4.096×10^9
$w = 6, u = 10, b = 3$	6	8.021×10^{56}

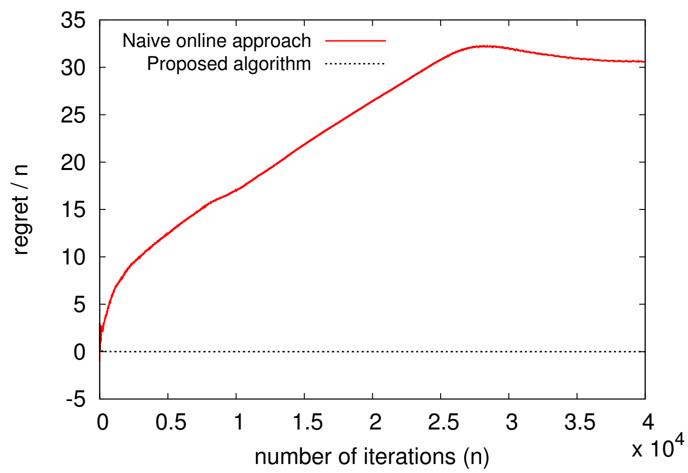


Fig. 10. Regret divided by the number of iterations for static **S** and **R**

represents the mean observed aggregate data dissemination success probability of the k-th data prefetching set, and m_k is the number of times that the k-th data prefetching set has been selected as the decision. As shown in Table II, the naive approach requires two data storages of size $\left(\frac{u!(u-b)!}{b!}\right)^w \times 1$ to store $\overline{Y}_k$ and m_k while the proposed algorithm requires two $w \times 1$ data storages. Therefore, the proposed online learning algorithm is more efficient than the naive approach in terms of memory resource usage.

Figure 10 depicts the simulation results of the regret divided by the number of iterations, where the vehicle route matrix

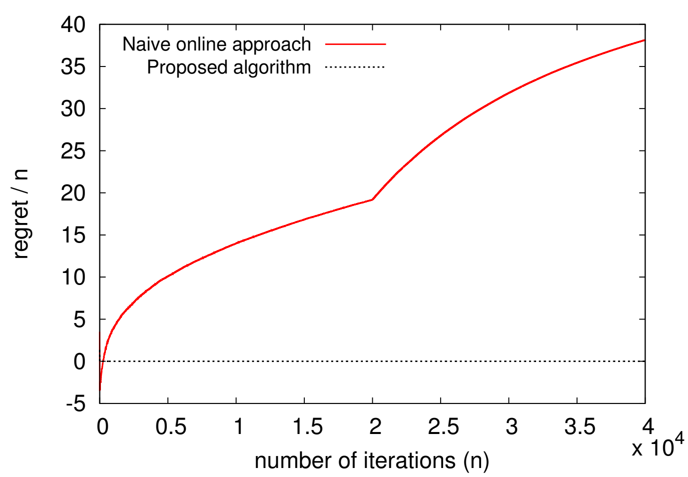


Fig. 11. Regret divided by the number of iterations where **R** is abruptly changed at the 20,000-th iteration.

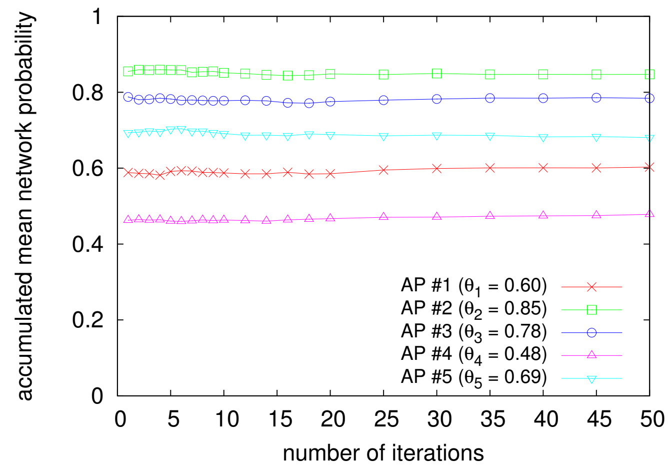


Fig. 12. Estimated network connectivity success rate of each roadside wireless AP with respect to iteration quantity.

S and the chunk request matrix **R** are unchanged during the simulation. Figure 10 shows that the regret of the proposed algorithm divided by the number of iterations converges to zero very rapidly in comparison with UCB1. It implies that the proposed online learning algorithm achieves better data dissemination performance than UCB1 in terms of convergence speed. Because the search space for UCB1 includes all possible data prefetching sets, UCB1 incurs a tremendous time cost to find the optimal data prefetching decision. Figure 11 depicts the regret divided by the number of iterations in a dynamically changing scenario where **R** changes at the 20,000-th iteration. As shown in Figure 11, the regret of the proposed algorithm converges consistently while that of the naive approach starts to increase after the 20,000-th iteration due to the change of **R**. This arises because UCB1 needs to re-learn all the aggregate dissemination success probabilities when the chunk requests change.

Figure 12 depicts the estimated network connectivity success rate of each AP during the iterations of the proposed online learning algorithm for the scenario in which **R** is changed for every iteration. Figure 12 shows that the estimated network connectivity success rate of the APs rapidly converges to their mean values even though **R** changes over time. This result implies that the proposed online learning algorithm is capable of finding the unknown network connectivity success rate of each AP in dynamically changing vehicular cloud service scenarios. The experiment results with real-life vehicle-traces and realistic wireless models indicate that the proposed dissemination method is applicable to data dissemination for VCSs.

VI. Conclusion

In this paper, we have presented a prefetch-based data dissemination method for VCSs consisting of roadside wireless APs with local data storages. We formulated the data dissemination problem as a combinatorial optimization that maximizes the aggregate data dissemination success probability when the size of local data storage is limited. Under the assumption

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2015.2388851, IEEE Transactions on Vehicular Technology

that the routes and data request information of vehicle subscribers are readily available, we devised two algorithms to determine how to prefetch a set of data from a data center to roadside wireless APs. The first is a greedy algorithm that solves the dissemination problem when wireless network connectivity characteristics are known deterministically. We proved that this algorithm could find a sub-optimal solution in a polynomial-time by deriving the approximation bound of the greedy algorithm. The second one is an MAB based online learning algorithm that gradually learns the unknown network connectivity success rate at each iteration and then determines an optimal binary decision matrix. In addition, we proved that its optimal solution converges to a globally optimal solution in a certain number of iterations using *regret* analysis. Finally, we presented numerical results of real-life vehicle-trace experiments to demonstrate the performance of the proposed algorithms in a variety of data dissemination scenarios in VCSs.

Appendix A
Approximation of the greedy algorithm

In this appendix, we show the worst-case performance of the greedy algorithm described in Algorithm 1 by deriving its approximation factor. The problem in (4) can be re-written as the following optimization problem:

$$\max_{\mathbf{X}\in\mathcal{F}} g_a(\mathcal{A}_{\mathbf{X}}) \text{ subject to } \mathbf{X}\mathbb{1}_u = b\cdot\mathbb{1}_w, \tag{17}$$

where $g_a(\mathcal{A}_{\mathbf{X}}) = \sum_{i=1}^{v}\sum_{j=1}^{u} r_{i,j}\prod_{k=1}^{w}(1-s_{i,k}\cdot\theta_k\cdot x_{k,j})$ is a nonnegative, nondecreasing, and submodular set function; thus, it satisfies the following condition:

$$g_a(T) \le g_a(S) + \sum_{i\in T\setminus S}(g_a(S\cup\{i\}) - g_a(S)), \tag{18}$$

where T and S are arbitrary sets. Assume that $\mathbf{X}^*$ is an optimal solution of the problem in (17), and $\mathbf{X}_g^i$ is a solution of the greedy algorithm after the i-th iteration. We sort the index set of $\{(m,n)\}_i$s for $\mathbf{X}^*$ such that

$$\begin{aligned} &g_a(\{(m,n)_1,\cdots,(m,n)_i\}) \\ &= \max_{(m,n)\in\mathcal{A}_{\mathbf{X}^*}\setminus\{(m,n)_1,\cdots,(m,n)_{i-1}\}} g_a(\{(m,n)_1,\cdots, \\ &\qquad (m,n)_{i-1}\}\cup\{(m,n)\}),\ i\in\{1,\cdots,w\cdot b\}. \end{aligned} \tag{19}$$

Let $Z=\{(m,n)_1,(m,n)_2,(m,n)_3\}$ be the set that consists of the first three elements of the index set $\mathcal{A}_{\mathbf{X}^*}$.[3] Then for any element $(m,n)_k\in\mathcal{A}_{\mathbf{X}^*}, k\ge 4$, and the set $Y\subset\mathcal{A}_{\mathbb{1}_{w\times u}}\setminus(Z\cup\{(m,n)_k\})$, the following inequalities hold:

$$\begin{aligned} &g_a(Z\cup Y\cup\{(m,n)_k\}) - g_a(Z\cup Y) \\ &\quad\le g_a(\{(m,n)_k\}) - g_a(\phi) \le g_a(\{(m,n)_1\}) \\ &g_a(Z\cup Y\cup\{(m,n)_k\}) - g_a(Z\cup Y) \\ &\quad\le g_a(\{(m,n)_1\}\cup\{(m,n)_k\}) - g_a(\{(m,n)_1\}) \\ &\quad\le g_a(\{(m,n)_1\}\cup\{(m,n)_2\}) - g_a(\{(m,n)_1\}) \\ &g_a(Z\cup Y\cup\{(m,n)_k\}) - g_a(Z\cup Y) \\ &\quad\le g_a(\{(m,n)_1\}\cup\{(m,n)_2\}\cup\{(m,n)_k\}) \\ &\qquad - g_a(\{(m,n)_1\}\cup\{(m,n)_2\}) \\ &\quad\le g_a(\{(m,n)_1\}\cup\{(m,n)_2\}\cup\{(m,n)_3\}) \\ &\qquad - g_a(\{(m,n)_1\}\cup\{(m,n)_2\}). \end{aligned}$$

By summing up all above inequalities, we obtain the following inequality:

$$g_a(Z\cup Y\cup\{(m,n)_k\}) - g_a(Z\cup Y) \le \frac{1}{3}g_a(Z). \tag{20}$$

Let $i^* = w\cdot b-3$ be the total number of iterations of the greedy algorithm and define a new function $f_a(S) = g_a(S) - g_a(Z)$, which is also a nondecreasing, nonnegative, and submodular set function if and only if $Z\subseteq S\subseteq\mathcal{A}_{\mathbb{1}_{w\times u}}$. According to the inequality (18), the following inequalities also hold:

$$\begin{aligned} &f_a(\mathcal{A}_{\mathbf{X}^*}) \\ &\le f_a(\mathcal{A}_{\mathbf{X}_g^i}) + \sum_{(m,n)\in\mathcal{A}_{\mathbf{X}^*}\setminus\mathcal{A}_{\mathbf{X}_g^i}}(f_a(\mathcal{A}_{\mathbf{X}_g^i}\cup\{(m,n)\}) - f_a(\mathcal{A}_{\mathbf{X}_g^i})) \\ &= f_a(\mathcal{A}_{\mathbf{X}_g^i}) + \sum_{(m,n)\in\mathcal{A}_{\mathbf{X}^*}\setminus\mathcal{A}_{\mathbf{X}_g^i}}(g_a(\mathcal{A}_{\mathbf{X}_g^i}\cup\{(m,n)\}) - g_a(\mathcal{A}_{\mathbf{X}_g^i})) \\ &\le f_a(\mathcal{A}_{\mathbf{X}_g^i}) + (w\cdot b-4)\times\vartheta_{i+1},\quad \forall i\in\{0,\cdots,i^*-1\}, \end{aligned} \tag{21}$$

where ϑ_i represents the maximum increment of the set function at the i-th iteration as follows:

$$\begin{aligned} \vartheta_i = &\max_{(m,n)\in\mathcal{A}_{\mathbb{1}_{w\times u}}\setminus(\mathcal{A}_{\mathbf{X}_g^{i-1}}\cup\mathcal{A}_{\mathbf{X}_c^{i-1}})} g_a(\mathcal{A}_{\mathbf{X}_g^{i-1}}\cup\{(m,n)\}) \\ &- g_a(\mathcal{A}_{\mathbf{X}_g^{i-1}}), \end{aligned} \tag{22}$$

where $\mathcal{A}_{\mathbf{X}_c^i} = \{(m,n)|x_{m,n}\notin\mathcal{A}_{\mathbf{X}^i}, (\mathbf{X}_g^{i-1}\cdot\mathbb{1}_u)_m \ge b,\ \forall m,n\}$ is an index set for the elements of which the value is 0 in $\mathbf{X}_g^i$ but at the same time are prohibited from being included in the index set $\mathcal{A}_{\mathbf{X}_g^i}$ due to the limited storage capacity of the roadside wireless APs. Thus, we obtain $g(\mathcal{A}_{\mathbf{X}_g^i}) = \sum_{j=1}^{i}\vartheta_i$ for all $i\in\{1,\cdots,i^*\}$.

To derive an approximation of the greedy algorithm, we use the inequality in [29]. If P and D are arbitrary positive integers, ρ_is are arbitrary nonnegative reals for $i=1,\cdots,P$, and $\rho_1>0$, then

$$\begin{aligned} \frac{\sum_{i=1}^{P}\rho_i}{\min_{t=1,\cdots,P}(\sum_{i=1}^{t-1}\rho_i + D\rho_t)} &\ge 1-\left(1-\frac{1}{D}\right)^P \\ &> 1-e^{P/D}. \end{aligned} \tag{23}$$

[3]Note that the greedy algorithm enumerates all feasible $\mathbf{X}\in\{\mathbf{X}:\mathbf{X}\in\{0,1\}^{w\times u}, |\mathcal{A}_{\mathbf{X}}|=3\}$, and then iteratively finds a sub-optimal solution starting from one of the feasible $\mathbf{X}$'s that gives the maximum of the cost function $g_a(\mathcal{A}_{\mathbf{X}})$. Thus, the first three elements of the set $Z\in\mathcal{A}_{\mathbf{X}^*}$ are also included in the index set of the greedy solution $\mathcal{A}_{\mathbf{X}_g^{i^*}}$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2015.2388851, IEEE Transactions on Vehicular Technology

Using (21) and (23), we obtain the following inequalities:

$$\begin{aligned}&\frac{f_a(\mathcal{A}_{\mathbf{X}_g^{i^*-1}} \cup \{(m,n)_{i^*}\})}{f_a(\mathcal{A}_{\mathbf{X}^*})}\\&\geq \frac{\sum_{j=1}^{i^*}\vartheta_i}{\min_{k=1,\cdots,i^*}\sum_{j=1}^{k-1}\vartheta_j + (w\cdot b-4)\times\vartheta_k}\\&\geq 1-e^{-1}.\end{aligned} \quad (24)$$

By combining (20) and (24),

$$\begin{aligned}&g_a(\mathcal{A}_{\mathbf{X}_g^{i^*}})\\&= g_a(Z) + f_a(\mathcal{A}_{\mathbf{X}_g^{i^*}})\\&\geq g_a(Z) + f_a(\mathcal{A}_{\mathbf{X}_g^{i^*-1}} \cup \{(m,n)_{i^*}\})\\&\quad - (f_a(\mathcal{A}_{\mathbf{X}_g^{i^*-1}} \cup \{(m,n)_{i^*}\}) - f_a(\mathcal{A}_{\mathbf{X}_g^{i^*-1}}))\\&= g_a(Z) + f_a(\mathcal{A}_{\mathbf{X}_g^{i^*-1}} \cup \{(m,n)_{i^*}\}) - (g_a(\mathcal{A}_{\mathbf{X}_g^{i^*-1}} \cup \{(m,n)_{i^*}\})\\&\quad - g_a(\mathcal{A}_{\mathbf{X}_g^{i^*-1}}))\\&\geq g_a(Z) + (1-e^{-1})f_a(\mathcal{A}_{\mathbf{X}^*}) - \frac{1}{3}g_a(Z)\\&= (e^{-1}-\frac{1}{3})g_a(Z) + (1-e^{-1})g_a(\mathcal{A}_{\mathbf{X}^*})\\&\geq (1-e^{-1})g_a(\mathcal{A}_{\mathbf{X}^*}).\end{aligned} \quad (25)$$

This demonstrates that the greedy algorithm achieves a constant factor approximation bounded by $(1-e^{-1})$.

APPENDIX B
UPPER BOUND OF THE COUNTER $T_k(N)$

In this appendix, we derive the upper bound of the counter $T_k(N)$. Let c_{n,m_k} denote $\sqrt{\frac{(w+1)\ln n}{m_k}}$ and I_n denote the index of the counter selected at the n-th iteration. Then, the upper bound of the counter $T_k(N)$ can be derived as the following inequalities:

$$\begin{aligned}&\mathbb{E}[T_k(N)] = 1 + \sum_{n=w+1}^{N} P\{I_n = k\}\\&\leq l + \sum_{n=w+1}^{N} P\{I_n = k, T_k(n-1)\geq l\}\\&\leq l + \sum_{n=w+1}^{N} P\Big\{\sum_{i=1}^{v}\sum_{j=1}^{u} r_{i,j}\Big(1-\prod_{k=1}^{w} h_{i,j,k}(s_{i,k},\overline{\theta}_k, x^*_{k,j})\\&\quad + 1 - \prod_{k=1}^{w}\max\{h_{i,j,k}(s_{i,k}, c_{n,m_k}, x^*_{k,j}),0\}\Big)\\&\quad \leq \sum_{i=1}^{v}\sum_{j=1}^{u} r_{i,j}\cdot\Big(1-\prod_{k=1}^{w} h_{i,j,k}(s_{i,k},\overline{\theta}_k, x_{k,j}(n))\\&\quad + 1 - \prod_{k=1}^{w}\max\{h_{i,j,k}(s_{i,k}, c_{n,m_k}, x_{k,j}(n)),0\}\Big),\\&\quad T_k(n-1)\geq l\Big\}\\&\leq l + \sum_{n=w+1}^{N} P\Big\{\min_{1<m_1,\cdots,m_w<t}\sum_{i=1}^{v}\sum_{j=1}^{u} r_{i,j}\\&\quad\times\Big(1-\prod_{k=1}^{w} h_{i,j,k}(s_{i,k},\overline{\theta}_k, x^*_{k,j}) + 1\\&\quad - \prod_{k=1}^{w}\max\{h_{i,j,k}(s_{i,k}, c_{n,m_k}, x^*_{k,j}),0\}\Big)\\&\quad\leq \max_{l\leq m_1,\cdots,m_w<t}\sum_{i=1}^{v}\sum_{j=1}^{u} r_{i,j}\\&\quad\times\Big(1-\prod_{k=1}^{w} h_{i,j,k}(s_{i,k},\overline{\theta}_k, x_{k,j}(n))\\&\quad + 1 - \prod_{k=1}^{w}\max\{h_{i,j,k}(s_{i,k}, c_{n,m_k}, x_{k,j}(n)),0\}\Big)\Big\}\\&\leq l + \sum_{n=1}^{\infty}\sum_{m_1=1}^{n-1}\cdots\sum_{m_w=1}^{n-1}\sum_{\overline{m}_1=l}^{n-1}\cdots\sum_{\overline{m}_w=l}^{n-1}\\&\quad P\Big\{\sum_{i=1}^{v}\sum_{j=1}^{u} r_{i,j}\cdot\Big(1-\prod_{k=1}^{w} h_{i,j,k}(s_{i,k},\overline{\theta}_k, x^*_{k,j})\\&\quad + 1 - \prod_{k=1}^{w}\max\{h_{i,j,k}(s_{i,k}, c_{n,m_k}, x^*_{k,j}),0\}\Big)\\&\quad\leq \sum_{i=1}^{v}\sum_{j=1}^{u} r_{i,j}\cdot\Big(1-\prod_{k=1}^{w} h_{i,j,k}(s_{i,k},\overline{\theta}_k, x_{k,j}(n))\\&\quad + 1 - \prod_{k=1}^{w}\max\{h_{i,j,k}(s_{i,k}, c_{n,\overline{m}_k}, x_{k,j}(n)),0\}\Big)\Big\},\end{aligned} \quad (26)$$

where l is an arbitrary positive integer. In order to satisfy the inequality in (26), at least one of the followings must hold:

$$\begin{aligned}&\sum_{i=1}^{v}\sum_{j=1}^{u} r_{i,j}\cdot(1-\prod_{k=1}^{w} h_{i,j,k}(s_{i,k},\overline{\theta}_k, x^*_{k,j}))\\&\quad\leq g(\mathbf{X}^*) - \sum_{i=1}^{v}\sum_{j=1}^{u} r_{i,j}\\&\quad\quad\times(1-\prod_{k=1}^{w}\max\{h_{i,j,k}(s_{i,k}, c_{n,m_k}, x^*_{k,j}),0\})\end{aligned} \quad (27)$$

$$\begin{aligned}&\sum_{i=1}^{v}\sum_{j=1}^{u} r_{i,j}\cdot(1-\prod_{k=1}^{w} h_{i,j,k}(s_{i,k},\overline{\theta}_k, x_{k,j}(n)))\\&\quad\geq g(\mathbf{X}(n)) + \sum_{i=1}^{v}\sum_{j=1}^{u} r_{i,j}\\&\quad\quad\times(1-\prod_{k=1}^{w}\max\{h_{i,j,k}(s_{i,k}, c_{n,\overline{m}_k}, x_{k,j}(n)),0\})\end{aligned} \quad (28)$$

$$\begin{aligned}&g(\mathbf{X}^*) < g(\mathbf{X}(n)) + 2\sum_{i=1}^{v}\sum_{j=1}^{u} r_{i,j}\\&\quad\times(1-\prod_{k=1}^{w}\max\{h_{i,j,k}(s_{i,k}, c_{n,\overline{m}_k}, x_{k,j}(n)),0\}).\end{aligned} \quad (29)$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2015.2388851, IEEE Transactions on Vehicular Technology

We next derive the upper bound of the probabilities for (27), (28) and (29). The upper bound for (27) is given by

$$
\begin{aligned}
&P\Big\{\sum_{i=1}^{v}\sum_{j=1}^{u} r_{i,j}\cdot(1-\prod_{k=1}^{w} h_{i,j,k}(s_{i,k},\overline{\theta}_k,x^*_{k,j}))\le g(\mathbf{X}^*)\\
&\quad-\sum_{i=1}^{v}\sum_{j=1}^{u} r_{i,j}\cdot(1-\prod_{k=1}^{w}\max\{h_{i,j,k}(s_{i,k},c_{n,m_k},x^*_{k,j}),0\})\Big\}\\
&\le\sum_{i=1}^{v}\sum_{j=1}^{u} r_{i,j}\cdot P\Big\{1-\prod_{k=1}^{w}\max\{h_{i,j,k}(s_{i,k},\overline{\theta}_k+c_{n,m_k},\\
&\qquad x^*_{k,j}),0\}\le 1-\prod_{k=1}^{w} h_{i,j,k}(s_{i,k},\theta_k,x^*_{k,j})\Big\}\\
&\le\sum_{i=1}^{v}\sum_{j=1}^{u} r_{i,j}\cdot(\sum_{k=1}^{w} P\{\overline{\theta}_k+c_{n,m_k}\le\theta_k\}+P\{1\le\theta_k\})\\
&\le\sum_{i=1}^{v}\sum_{j=1}^{u} r_{i,j}\cdot(\sum_{k=1}^{w} P\{\overline{\theta}_k+c_{n,m_k}\le\theta_k\}).
\end{aligned}\tag{30}
$$

We use the Chernoff-Hoeffding bound in (30). Let $X_1,\cdots,X_n$ be random variables in the range $[0,1]$ such that $\mathbb{E}[X_t|X_1,\cdots,X_{t-1}]=\mu$, and $S_n=X_1+\cdots+X_n$. Then, for all $a\ge 0$

$$
P\{S_n\ge n\mu+a\}\le e^{-2a/n}\text{ and }P\{S_n\ge n\mu-a\}\le e^{-2a/n}.\tag{31}
$$

By the above inequality, the upper bound of the probability $P\{\overline{\theta}_k+c_{n,m_k}\le\theta_k\}$ for all $k=\{1,\cdots,w\}$ is given by

$$
\begin{aligned}
P\{\overline{\theta}_k+c_{n,m_k}\le\theta_k\}&\le e^{-2m_k^2\cdot\left(\sqrt{\frac{(w+1)\ln n}{m_k}}\right)^2\cdot\frac{1}{m_k}}\\
&=e^{-2(w+1)\ln n}=n^{-2(w+1)}.
\end{aligned}\tag{32}
$$

Similarly, the upper bound of the probability for (28) is also given by

$$
\begin{aligned}
P\{\overline{\theta}_k\ge\theta_k+c_{n,\overline{m}_k}\}&\le e^{-2\overline{m}_k^2\cdot\left(\sqrt{\frac{(w+1)\ln n}{\overline{m}_k}}\right)^2\cdot\frac{1}{\overline{m}_k}}\\
&=e^{-2(w+1)\ln n}=n^{-2(w+1)}.
\end{aligned}\tag{33}
$$

Last, we consider the last inequality in (29). Because the dissemination failure probability $p_{i,j}(\boldsymbol{\theta})$ is a decreasing function with respect to $\boldsymbol{\theta}$ for all i and j, there exists $\boldsymbol{\theta}_{ij}$ for the i-j pair such that the following inequality holds:

$$
p_{i,j}=\prod_{k=1}^{w} h_{i,j,k}(s_{i,k},(\boldsymbol{\theta}_{ij})_k,x_{k,j})\ge 1-\frac{\delta_{\min}}{2wb},\tag{34}
$$

where $\delta_{\min}=g(\mathbf{X}^*)-\max_{\mathbf{X}\in\mathcal{F}\setminus\mathbf{X}^*}g(\mathbf{X})$. By (34), the following inequalities hold:

$$
\begin{aligned}
&\sum_{i=1}^{v}\sum_{j=1}^{u} r_{i,j}\cdot(1-\prod_{k=1}^{w} h_{i,j,k}(s_{i,k},\theta_{\min},x_{k,j}),x_{k,j}))\\
&\le\sum_{i=1}^{v}\sum_{j=1}^{u} r_{i,j}\cdot p_{i,j}(\boldsymbol{\theta}_{ij})\le\frac{\delta_{\min}}{2},
\end{aligned}\tag{35}
$$

where $\theta_{\min}=\min_{(i,j,k)\in\{(i,j,k)|1\le i\le v,1\le j\le u,1\le k\le w\}}(\boldsymbol{\theta}_{ij})_k$. If we choose the integer $l\ge\left\lceil\frac{(w+1)\ln N}{\theta^2_{min}}\right\rceil$ in (26), the inequality in (29) does not hold:

$$
\begin{aligned}
&g(\mathbf{X}^*)-g(\mathbf{X}(n))-2\sum_{i=1}^{v}\sum_{j=1}^{u} r_{i,j}\\
&\quad\times(1-\prod_{k=1}^{w}\max\{h_{i,j,k}(s_{i,k},c_{n,\overline{m}_k},x_{k,j}(n)),0\})\\
&=g(\mathbf{X}^*)-g(\mathbf{X}(n))-2\sum_{i=1}^{v}\sum_{j=1}^{u} r_{i,j}\\
&\quad\times(1-\prod_{k=1}^{w}\max\{h_{i,j,k}(s_{i,k},\sqrt{\frac{(w+1)\ln N}{\overline{m}_k}},x_{k,j}(n)),0\})\\
&\ge g(\mathbf{X}^*)-g(\mathbf{X}(n))-2\sum_{i=1}^{v}\sum_{j=1}^{u} r_{i,j}\\
&\quad\times(1-\prod_{k=1}^{w}\max\{h_{i,j,k}(s_{i,k},\sqrt{\frac{(w+1)\ln N}{l}},x_{k,j}(n)),0\})\\
&\ge g(\mathbf{X}^*)-g(\mathbf{X}(n))-2\sum_{i=1}^{v}\sum_{j=1}^{u} r_{i,j}\\
&\quad\times(1-\prod_{k=1}^{w}\max\{h_{i,j,k}(s_{i,k},\theta_{min},x_{k,j}(n)),0\})\\
&\ge g(\mathbf{X}^*)-g(\mathbf{X}(n))-\delta_{\min}\ge 0.
\end{aligned}
$$

Therefore, we do not need to consider the last case in (29). Based on the above inequalities, the upper bound of the counter $\mathbb{E}[T_k(N)]$ is given as follows:

$$
\begin{aligned}
&\mathbb{E}[T_k(N)]\\
&\le l+\sum_{n=2}^{\infty}\sum_{m_1=1}^{n-1}\cdots\sum_{m_w=1}^{n-1}\sum_{\overline{m}_1=l}^{n-1}\cdots\sum_{\overline{m}_w=l}^{n-1}\Big\{\sum_{i=1}^{v}\sum_{j=1}^{u} r_{i,j}\\
&\qquad\times(\sum_{k=1}^{w} P\{\overline{\theta}_k+c_{n,m_k}\le\theta_k\}+P\{\overline{\theta}_k\le\theta_k+c_{n,\overline{m}_k}\})\Big\}\\
&\le\left\lceil\frac{(w+1)\ln N}{\theta^2_{\min}}\right\rceil+\sum_{n=2}^{\infty}\sum_{m_1=1}^{n-1}\cdots\sum_{m_w=1}^{n-1}\sum_{\overline{m}_1=l}^{n-1}\cdots\sum_{\overline{m}_w=l}^{n-1}2\\
&\qquad\times(w^2\cdot b)n^{-2(w+1)}\\
&\le\frac{(w+1)\ln N}{\theta^2_{\min}}+1+(w^2\cdot b)\sum_{n=1}^{\infty}2n^{-2}\\
&\le\frac{(w+1)\ln N}{\theta^2_{\min}}+1+\frac{\pi^2}{3}(w^2\cdot b).
\end{aligned}\tag{36}
$$

References

[1] H. Hartenstein and K. P. Laberteaux, "A tutorial survey on vehicular ad hoc networks," in *IEEE Communications Magazine*, vol. 46, no. 6, pp. 164-171, 2008.

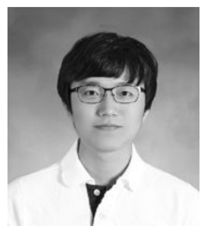
[2] M. J. Khabbaz, C. M. Assi, and W. F. Fawaz, "Disruption-tolerant networking: a comprehensive survey on recent developments and persisting challenges," in *IEEE Communications Surveys and Tutorials*, vol. 14, no. 2, pp. 607-640, 2012.

[3] M. Whaiduzzaman, M. Sookhak, A. Gani, and R. Buyya, "A survey on vehicular cloud computing," in *Journal of Network and Computer Applications*, vol. 40, no. 1, pp. 325-344, 2014.

[4] P. Bellavista, A. Corradi, M. Fanelli, and L. Foschini, "A survey of context data distribution for mobile ubiquitous systems," in *ACM Computing Surveys*, vol. 44, no. 4, pp. 1-49, 2013.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2015.2388851, IEEE Transactions on Vehicular Technology

[5] J. Zhao, Y. Zhang, and G. Cao, "Data pouring and buffering on the road: a new data dissemination paradigm for vehicular ad hoc networks," in *IEEE Transactions on Vehicular Technology*, vol. 56, no. 6, pp. 3266-3277, 2007.
[6] B. B. Dubey, N. Chauhan, and S. Pant, "RRDD: reliable route based data dissemination technique in vanets," in *Proceedings of the 2011 International Conference on Communication Systems and Network Technologies (CSNT)*, 2011.
[7] X. Bai, M. Gong, Z. Gao, and S. Li, "Reliable and efficient data dissemination protocol in vanets," in *Proceedings of the IEEE 8th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM)*, 2012.
[8] R. S. Schwarts, A. E. Ohazulike, C. Sommer, H. Scholten, F. Dressler, and P. Havinga, "Fair and adaptive data dissemination for traffic information systems," in *Proceedings of the 2012 IEEE Vehicular Networking Conference (VCN)*, 2012.
[9] F. Ye, S. Roy, and H. Wang, "Efficient data dissemination in vehicular ad hoc networks," in *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 4, pp. 769-779, 2012.
[10] M. Sathiamoorthy, A. G. Dimakis, B. Krishnamachari, and F. Bai, "Distributed storage codes reduce latency in vehicular networks," in *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM)*, 2012.
[11] H. Liang and W. Zhuang, "Cooperative data dissemination via roadside WLANs," in *IEEE Communications Magazine*, vol. 50, no. 4, pp. 68-74, 2012.
[12] M. A. Salkuyeh, F. Hendessi, and T. A. Gulliver, "Data dissemination with rateless coding in a grid vehicular topology," in *EURASIP Journal on Wireless Communications and Networking*, vol. 2013, no. 1, pp. 1-14, 2013.
[13] Y. Ding and L. Xiao, "SAVD: static-node-assisted adaptive data dissemination in vehicular networks," in *IEEE Transactions on Vehicular Technology*, vol. 59, no. 5, pp. 2445-2455, 2010.
[14] J. Ott and D. Kutscher, "Drive-thru internet: IEEE 802.11b for "automobile" users," in *Proceedings of the Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, 2004.
[15] V. Bychkovsky, B. Hull, A. Miu, H. Balakrishnan, and S. Madden, "A measurement study of vehicular internet access using in situ wifi networks," in *Proceedings of the ACM 12th Annual International Conference on Mobile Computing and Networking (MobiCom)*, 2006.
[16] J. Eriksson, H. Balakrishnan, and S. Madden, "Cabernet: vehicular content delivery using wifi," in *Proceedings of the ACM 14th Annual International Conference on Mobile Computing and Networking (MobiCom)*, 2008.
[17] J. Jeong, S. Guo, Y. Gu, T. He, and D. H. Du, "Trajectory-based statistical forwarding for multihop infrastructure-to-vehicle data delivery," in *IEEE Transactions on Mobile Computing*, vol. 11, no. 10, pp. 1523-1537, 2012.
[18] J. Wang, "A survey of web caching schemes for the internet," in *ACM SIGCOMM Computer Communication Review*, vol. 29, no. 5, pp. 36-46, 1999.
[19] B. Li, M. J. Golin, G. F. Italiano, and X. Deng, "On the optimal placement of web proxies in the internet," in *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM)*, 1999.
[20] L. Qui, V. J. Padmanabhan, and G. M. Voelker, "On the placement of web server replicas," in *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM)*, 2001.
[21] A. Nimkar, C. Mandal, and C. Reade, "Video placement and disk load balancing algorithm for VoD proxy server," in *Proceedings of the IEEE International Conference on Internet Multimedia Services Architecture and Applications (IMSAA)*, 2009.
[22] I. Giurgiu, C. Castillo, A. Tantawi, and M. Steinder, "Enabling efficient placement of virtual infrastructures in the cloud," in *Proceedings of the International Middleware Conference (Middleware)*, 2012.
[23] Y. Gao, H. Guan, Z. Qi, Y. Hou, and L. Liu, "A multi-objective ant colony system algorithm for virtual machine placement in cloud computing," in *Journal of Computer and System Sciences*, vol. 79, no. 8, pp. 1230-1242, 2013.
[24] K. Son, H. Kim, Y. Yi, and B. Krishnamachari, "Base station operation and user association mechanisms for energy-delay tradeoffs in green cellular networks," in *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 8, pp. 1525-1536, 2011.
[25] Y. Gai and B. Krishnamachari, "Online learning algorithms for stochastic water-filling," in *proceedings of Information Theory and Applications Workshop (ITA)*, 2012.
[26] P. Auer, N. Cesa-Bianchi, and P. Fischer, "Finite-time analysis of the multiarmed bandit problem," in *Machine Learning*, vol. 47, no. 2-3, pp. 235-256, 2002.
[27] Y. Gai, B. Krishnamachari, and R. Jain, "Combinatorial network optimization with unknown variables: multi-armed bandits with linear rewards and individual observations," in *IEEE/ACM Transactions on Networking*, vol. 20, no. 5, pp. 1466-1478, 2012.
[28] M. Sathiamoorthy, A. G. Dimakis, B. Krishnamachari, and F. Bai, "Distributed storage codes reduce latency in vehicular networks," in *Proceedings of the Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, 2012.
[29] L. A. Wolsey, "Maximizing real-valued submodular functions: primal and dual heuristics for location problems," in *Mathematics of Operations Research*, vol. 7, no. 3, pp. 410-425, 1982.



Ryangsoo Kim received his B.S. degree in Information and Communications from Chungnam National University, Daejeon, Republic of Korea, in 2010, and his M.S. degree in Information and Communications from the Gwangju Institute of Science and Technology (GIST), Republic of Korea, in 2012. He is currently pursuing a Ph.D. degree in at the School of Information and Communications at GIST. His research interests include location-aware applications in wireless networks and data dissemination in VANET.

Hyuk Lim (M'03) received his B.S., M.S., and Ph.D. degrees all from the School of Electrical Engineering and Computer Science, Seoul National University (SNU), Seoul, Korea in 1996, 1998, and 2003, respectively. He is currently an associate professor in the School of Information and Communications, Gwangju Institute of Science and Technology (GIST), Gwangju, Republic of Korea. He was a postdoctoral research associate in the Department of Computer Science, University of Illinois at Urbana-Champaign in 2003-2006. His research interests include the analytical modeling and empirical evaluation of computer networking systems, and the network protocol design and performance analysis for wireless networks.

Bhaskar Krishnamachari (M'03) received his B.E. in Electrical Engineering at The Cooper Union, New York, in 1998, and his M.S. and Ph.D. degrees from Cornell University in 1999 and 2002 respectively. He is currently an Associate Professor and a Ming Hsieh Faculty Fellow in the Department of Electrical Engineering at the University of Southern California's Viterbi School of Engineering. His primary research interest is in the design and analysis of algorithms and protocols for next-generation wireless networks.