DOVE: Data Offloading through Spatio-temporal Rendezvous in Vehicular Networks

Munyoung Lee*, Junghwan Song*, Jaehoon (Paul) Jeong†, and Ted "Taekyoung" Kwon*
*School of Computer Science and Engineering, Seoul National University, Seoul 151-742, Republic of Korea
†Department of Interaction Science, Sungkyunkwan University, Suwon, Gyeonggi-Do 440-746, Republic of Korea
Emails: {mylee, jhsong}@mmlab.snu.ac.kr, pauljeong@skku.edu, tkkwon@snu.ac.kr

***Abstract*—The increasing mobile traffic is becoming a serious concern for mobile network providers. To address the traffic explosion problem, there have been a lot of efforts to offload the traffic from cellular networks to other networks, such as WiFi hotspots and femtocells. In this paper, we explore the potential benefits of vehicular networks for data offloading and propose a Data Offloading framework using Vehicular nEtworks (DOVE), which reduces the cellular traffic for in-vehicle data services in a cost effective way. DOVE exploits vehicle trajectories for offloading purposes so that content files requested by vehicles can be delivered via vehicular networks rather than cellular networks for economical purposes. We formulate the problem of selecting offloading positions as a spatio-temporal set-covering problem, and propose a time-prediction based set-covering algorithm using vehicle trajectories. Simulation results show that our DOVE framework can significantly reduce 57% of cellular link usage by performing data offloading through vehicular networks.**

Keywords*—*Data Offloading, Vehicular Network, Trajectory, Content Sharing, Set-Covering Problem, Road Network.

I. Introduction

Vehicular networks have emerged as one of the promising research areas to support the driving safety such as vehicle collision avoidance. To increase the driving safety and support reliable communications among vehicles, IEEE has standardized Dedicated Short Range Communications (DSRC) under the name of IEEE 802.11p [1], and static infrastructure nodes (i.e., relay nodes (RNs) and road-side units (RSUs)) are deployed at intersections to cope with the interference by road environments (e.g., buildings) in the wireless communications for the driving safety [2], [3]. Also, automotive vendors have plans to embed cellular communication modules into cars to enable in-vehicle data services [4]. Through the cellular communications, drivers and passengers can be provided with various in-vehicle data services along with the driving safety. For instance, users can enjoy *delay-insensitive* in-vehicle data services, such as software update for car system and rich site summary (RSS) services (as non-real time applications) for car dash screen (e.g., news headlines and audio/video clips for entertainment). These delay-insensitive services can be available to drivers and passengers via non-cellular links with some delay to reduce the cost of cellular link usage.

Considering the current trend of traffic explosion in mobile environments [5], if in-vehicle data services using cellular networks become prevalent, it will significantly worsen the problem of cellular traffic explosion [6]. To address this traffic explosion problem not only in vehicular environments, but also in general settings, there have been a lot of efforts to offload the traffic from cellular networks to other networks, such as WiFi hotspots and femtocells [7]–[9]. Also, recent measurement study [10] shows that the requests of some popular videos account for the majority of all the requests, and a significant amount of cellular traffic is redundant. Thus, we focus on the data offloading for redundant traffic caused by in-vehicle data services in order to mitigate the traffic overload in cellular networks. To this end, a *natural research question* is how to design an efficient offloading framework utilizing components of vehicular networks (i.e., relay nodes) to provide the delay-insensitive in-vehicle data services, while minimizing the usage of expensive cellular links.

To design an offloading framework leveraging vehicular networks, we focus on the following characteristics of vehicular networks. First, a vehicle is equipped with multiple communication interfaces (e.g., DSRC and 4G-LTE), a GPS navigation system, and a storage module. Thus, vehicles can carry volume-intensive files like multimedia on their storage and communicate with each other by DSRC communications. Second, the vehicle mobility is predictable since (i) vehicles are moving along the constrained roadways and (ii) their travel paths (i.e., navigation paths) to destinations can be calculated from navigation systems. Last, infrastructure nodes (e.g., RNs [2], [3]) that are deployed in vehicular networks for the driving safety [11] can be used for data offloading. With the mobility information and infrastructure nodes of vehicular networks that are available for data offloading, a significant amount of the cellular traffic for vehicles can be reduced.

In this paper, we propose a Data Offloading framework through Vehicular nEtworks (DOVE) that can significantly reduce the cellular traffic for delay-insensitive in-vehicle data services in a cost-effective way. The key idea in DOVE is to predict vehicle mobility with vehicle trajectories provided by the GPS navigation systems. With this vehicle mobility prediction, DOVE selects appropriate *offloading positions (OPs)* where vehicles can retrieve the requested files without using cellular links. Note that RNs play the role of OPs since RNs can keep and deliver files to the passing vehicles that request the files. In this paper, we formulate the selection of OPs (i.e., RNs) as a spatio-temporal set-covering problem where each offloading position covers a set of vehicles that request the same file. To solve the selection of OPs, we propose a *time-prediction based set-covering algorithm* (called *DOVE algorithm*) using vehicle trajectories. This *DOVE algorithm* selects approximately the minimum number of OPs, reducing a significant amount of aggregated usage of cellular links for vehicles. Our contributions in this paper are as follows:

- **DOVE**: we propose a data offloading framework (called *DOVE*) using the components of vehicular

978-1-4799-9964-4/15/$31.00 ©2015 IEEE

networks (i.e., RNs) for delay-insensitive in-vehicle data services.
- **DOVE algorithm**: we formulate the selection of offloading positions as a spatio-temporal set-covering problem. To solve this problem, we propose a time-prediction based set-covering algorithm (called *DOVE algorithm*) using vehicle trajectories to select offloading positions.

The rest of this paper is organized as follows. Section II describes a problem formulation. Section III explains the design and operations of DOVE framework. In Section IV, we evaluate the performance of our DOVE algorithm. Section V summarizes related work. Finally, this paper is concluded along with future work in Section VI.

II. Problem Formulation

A. Target Scenario and Goal

A significant amount of cellular traffic is redundant since multiple users repeatedly download the same popular file [10]. Accordingly, a substantial reduction of cellular traffic for vehicles is expected by offloading the duplicated traffic (of popular files) to vehicular networks. Thus, *our target* is to reduce the cellular traffic of popular files for delay-insensitive in-vehicle data services. Target content files can be (i) update files for software in car system and (ii) popular multimedia files (e.g., headline news, music files, and YouTube video clips).

If these content files are delay-insensitive, it can be assumed that users are willing to wait for some delay to reduce the cost of cellular links [9]. In other words, users in vehicles face a trade-off between cost and delay in making their offloading decisions. In our scenario, vehicles try to offload their traffic from cellular links to RNs in vehicular networks when users choose to wait in order to reduce their cellular service cost.

In our target scenario, *our goal* is to select effective offloading positions (OPs), such as RNs that minimize the aggregated usage of cellular links for vehicles (i.e., the amount of data downloaded through cellular networks), while satisfying the user-defined *quality of experience (QoE)* requirements in terms of content retrieval delay.

B. DOVE Components and Assumptions

As depicted in Fig. 1, we describe the components of vehicular offloading framework (DOVE) with assumptions.

◇ **A traffic control center (TCC)** [2], [12] is a traffic management node that maintains vehicle trajectories for the location management of vehicles. In DOVE, the TCC collects the requested content information from vehicles. Thus, it knows the set of vehicles (called *request vehicles*) that request the same content file. Note that vehicles can communicate with the TCC via cellular links, whose traffic volume is negligible. Also, multiple TCCs can be used for the scalability and survivability by dividing the road network into several regions.

◇ **A relay node (RN)** is a wireless packet holder for the reliable forwarding, which is usually deployed in vehicular networks for the driving safety [2], [11]. Compared with a road-side unit (RSU) [13], an RN also has the DSRC communications and storage modules, but does not have the Internet

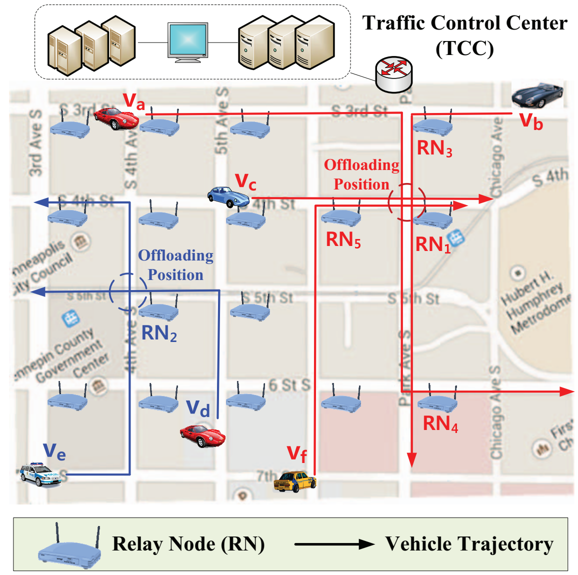

Fig. 1. DOVE: Data offloading framework using RNs in vehicular networks.

connectivity for the cost effectiveness. Note that the cost for installing an RSU is as much expensive as US $5,000 [14] due to the installation of wireline for Internet connectivity. Since RNs can keep and share content files using their storage, we can exploit RNs for data offloading (role of OPs) as well as driving safety. It is assumed that one RN is deployed at each intersection for the driving safety [11]. However, DOVE works even in the case where some intersections do not have their own RNs. The details will be described in Section IV-D.

◇ **Vehicles** participating in vehicular networks have DSRC devices and cellular communication devices [1], [4]. Thus, vehicles can communicate with RNs using DSRC devices to retrieve the requested file.

◇ Vehicles, the TCC, and the RNs are equipped with GPS navigators and digital road maps. Recent commercial navigation systems provide vehicles with the traffic statistics information, such as average vehicle speed v and vehicle arrival rate λ per road segment [15].

◇ When users in vehicles want to retrieve a content file, they should make a decision: (i) consuming cellular links, immediately, with cellular link cost, or (ii) waiting for offloading via vehicular networks, within tolerance time δ, without cellular link cost. If users decide to use delay-insensitive content retrieval through data offloading, they send their current location, trajectory, and content request to the TCC via cellular links. Note that TCC can perform the data offloading schedule with the vehicle trajectories. Even though all of the drivers are not using GPS navigators for their everyday commute, we assume that the drivers participating in DOVE service are using GPS navigators. For the vehicles without GPS navigator, we do not count them as DOVE users.

C. Design Principles using RNs

DOVE is designed to use RNs for the role of OPs due to the following reasons: (i) *High success probability* and (ii) *Cost effectiveness*. First, to use vehicles as OPs, vehicles requesting the same file should encounter each other and establish connectivity within the DSRC range. However, the probability of such occasions will be very small. On the other hand, when we use static infrastructure nodes (i.e., RNs) as OPs, the file can be shared with any vehicles that will pass the RNs. The vehicles reaching RNs with the file can get data from the RNs at any time. Thus, the data offloading using these RNs can provide reliable data delivery for moving vehicles. Second, the cost for installing an RN is relatively cheaper than that of an RSU [2], [14]. The reason is that as stand-alone wireless nodes, RNs do not have the Internet connectivity unlike RSUs.

D. The Concept of Offloading in DOVE

RNs are infrastructure nodes usually deployed at intersections for the driving safety [11]. In DOVE, RNs play the role of OPs for offloading purposes. Given the request vehicles and their trajectories, we can find RNs where trajectories are overlapped. We call them the candidates for OPs and we will select appropriate RNs among the candidates. In Fig. 1, RN_1-RN_5 are candidates for OPs. If a selected RN has the requested file, vehicles passing through the selected RN can retrieve the file by DSRC communications. Since the content file is retrieved from the vehicular network, vehicles can reduce the cost of cellular link usage. In this case, the selected RN serves as a *spatio-temporal rendezvous* for request vehicles. Thus, vehicles can offload their traffic from the cellular links to the DSRC links to RNs in the vehicular network.

To use the selected RN as an OP, the RN should hold the file for *request vehicles* passing through the RN. To make the requested file be located in the selected RN, a vehicle that will arrive first at the selected RN is scheduled to download the file using cellular links. Also, it will store the downloaded file into the selected RN for other *request vehicles*. We call this vehicle a *provider*. On the other hand, other *request vehicles* except the *provider* are scheduled to retrieve the requested file from the selected RN without using cellular links. We call these vehicles *consumers*. In Fig. 1, RN_1 is selected as the OP and vehicle v_c is a provider. Thus, the provider v_c stores the file in RN_1 and consumers v_a, v_b, and v_f can retrieve it via the vehicular network. In summary, when an RN is selected as an OP, the RN receives the content file from the provider, keeping the file in its local storage. The stored content file can then be delivered by the RN to other vehicles (i.e., consumers) when they reach the communication range of the RN.

III. Design and Operations of DOVE

In this section, we explain the design and operations of our DOVE framework. First, we model the travel time of vehicle. Second, we explain the design and operations of our DOVE through a selection algorithm for offloading positions along with the selection of providers as follows: (i) the collection of content requests of vehicles by TCC, (ii) the selection of OPs, (iii) the selection of providers, and (iv) the selection of comsumers.

A. Travel Time Prediction

We model the travel time of a vehicle from one position to another position in a given road network. Using the travel time prediction and vehicle trajectories, the arrival time of vehicle at a particular RN can be calculated.

*1) **Travel Time through Road Segment**:* The travel time of vehicle over a fixed distance follows the Gamma distribution in light-traffic road condition [2], [16]. Thus, the travel time through a road segment i in the road network (called *link travel time*) is modeled as: $d_i \sim \Gamma(\kappa_i, \theta_i)$ where κ_i is a shape parameter and θ_i is a scale parameter. The parameters κ_i and θ_i are computed with the mean travel time μ_i and the travel time variance σ_i^2 [2]. Note that the traffic statistics of μ_i and σ_i^2 can be computed by commercial navigation service provider [15].

Let the mean of d_i be $E[d_i] = \mu_i$ and the variance of d_i be $Var[d_i] = \sigma_i^2$, the formulas for κ_i and θ_i are as follows:

$$\theta_i = \frac{Var[d_i]}{E[d_i]} = \frac{\sigma_i^2}{\mu_i} \tag{1}$$

$$\kappa_i = \frac{E[d_i]}{\theta_i} = \frac{\mu_i^2}{\sigma_i^2} \tag{2}$$

*2) **Travel Time on End-to-End Path** :* As described above, the travel time through a road segment i is modeled as the Gamma distribution of $d_i \sim \Gamma(\kappa_i, \theta_i)$. Given a vehicle trajectory, we assume that the travel times of road segments consisting of the trajectory are independent. Under this assumption, we approximate the mean and variance of End-to-End (E2E) travel delay as the sum of the means and the sum of the variances of the link travel times along the trajectory, respectively. Assuming that the traveling path consists of N road segments, the mean and variance of the E2E travel delay D are computed as follows:

$$E[D] = \sum_{i=1}^{N} E[d_i] = \sum_{i=1}^{N} \mu_i \tag{3}$$

$$Var[D] = \sum_{i=1}^{N} Var[d_i] = \sum_{i=1}^{N} \sigma_i^2 \tag{4}$$

With (3) and (4), the E2E vehicle delay distribution can be modeled as a Gamma distribution as follows: $D \sim \Gamma(\kappa_D, \theta_D)$ such that κ_D and θ_D are calculated using E[D] and Var[D] using the formulas of (1) and (2). Note that we can use any better E2E travel time distribution if it is available from actual measurement or mathematical model.

Using the travel time prediction, we can estimate vehicle's arrival time at a particular RN as follows: Let T^* be the current time. Let $t_{a,b}$ be the E2E travel time from the current position a to the RN b. We define $T_{v,p_v i}$ as the vehicle v's arrival time at RN i from the current position p_v. Then, the arrival time can be modeled as a Gamma distribution with Equations (3) and (4) such that $T_{v,p_v i}$ = T^* + $t_{p_v,i}$.

B. The Operation of TCC

To decide which RNs are used as OPs, the TCC collects the content request and the trajectory from each vehicle. The TCC constructs a set of *request vehicles* with vehicles requesting the same file. Also, it finds the candidates for OPs as RNs through which the trajectories of the vehicles pass. Next, the TCC tries to select the optimal OPs among the OP candidates and the providers for the selected OPs by the *selection algorithm*, described in Section III-C. Finally, the TCC sends the offloading information to each vehicle. The

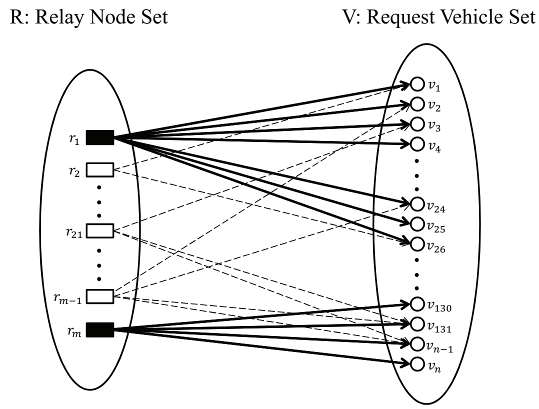

Fig. 2. The set-cover of RN set for request vehicle set.

offloading information includes the information of an OP and the role of a vehicle (i.e., *provider* or *consumer*). If there is no available OP, vehicles receive this information from the TCC and use cellular links to download the file.

C. The Selection Algorithm for Offloading Positions

In this subsection, we formulate the selection of OPs as a spatio-temporal set-covering problem, and propose a time-prediction based set-covering algorithm (called *DOVE algorithm*). Selecting an effective OP is important because the reduced cellular traffic volume and the content retrieval delay are determined by the OP. To reduce the aggregated usage of cellular links for vehicles, the number of providers using cellular links should be minimized. Since each provider consumes cellular traffic to provide the requested content file with an appropriate OP, the reduction of OPs can lower the number of providers. Therefore, we try to minimize the aggregated cellular traffic by selecting the minimum number of OPs. In other words, *the problem of minimizing cellular traffic can be formulated as the problem of selecting the minimum number of OPs.*

Let R be a set of RNs in the road network and V be a set of request vehicles requesting the same content file. Then, the selection of OPs can be formulated as the relation between set R and set V. In Fig. 2, a line (both solid and dashed) between an RN and a vehicle means that the vehicle will traverse the RN and hence can retrieve the file from the RN. If we select r_1 and r_m as OPs, they can *cover* all the request vehicles (represented by solid lines). This relation can be formulated as a *set-covering problem* that is one of NP-hard problems.

Let V be the set of request vehicles, and let S_i be the set of request vehicles covered by an RN i. Let F be a family of subsets of V, that is, $S_i \subseteq V$ such that $V = \bigcup_{S_i \in F} S_i$ where $F = \{S_i | i \in R\}$. Also, we define the collection of subsets $C \subseteq F$ as the set-cover of V such that $V = \bigcup_{S_i \in C} S_i$. Then, our set-covering problem is to find a minimum set-cover C^* of RNs as OPs as follows:

$$C^* \leftarrow \arg \min_{C \subseteq F} |C|, \tag{5}$$

where $V = \bigcup_{S_i \in C} S_i$. Let the set P be the set of offloading pairs, which consist of an RN and a provider. The identifier i of the element sets $S_i \in C$ means the selected RN i as an OP, which covers the request vehicles in S_i. Next, we propose a time-prediction based set-covering algorithm (called *DOVE algorithm*). Also, we propose an enhanced version of *DOVE algorithm* to further reduce the usage of cellular links.

1) DOVE Algorithm: We design the DOVE algorithm based on the greedy approach because it is known as the best possible polynomial time approximation algorithm for the set-covering algorithm under reasonable complexity assumptions [17]. Given a set of RNs R and a set of request vehicles V, the idea is to select the set S_i covering the largest number of the remaining vehicles not covered yet in each step in order to select the next RN as an OP, considering the travel times of vehicles.

Algorithm 1 DOVE Algorithm (R, V, F)

$I \leftarrow R$
$U \leftarrow V$
$P \leftarrow \emptyset$
while $U \neq \emptyset$ **do**
  update $S_i^* \leftarrow S_i$ for $i \in I$ by **pruning** unsatisfied vehicles v such that $\hat{t}_{p_v,i} < \gamma$ or $t_{p_v,i} > \delta$ where $v \in S_i$.
  select a $S_i^* \in F$ that maximizes $|S_i^* \cap U|$ for $i \in I$
  select a provider $d_i \in S_i^*$ whose arrival time at RN i is minimum
  $U \leftarrow U - S_i^*$
  $I \leftarrow I - \{i\}$
  $P \leftarrow P \cup \{(i, d_i)\}$
end while
return P

Let $\hat{t}_{p_v,i}$ be the travel time of vehicle v from the current position p_v to out of the communication range of an OP i; note that in this case, the vehicle v is within the communication range of the OP i. Let $t_{p_v,i}$ be the travel time of v from the current position p_v to an OP i, as defined in Section III-A2. We use γ to denote the sum of cellular download time and DSRC upload time when a provider provides a file for an RN via cellular and DSRC links. The γ value is decided by the *size of file* and the *bandwidthes of cellular and DSRC links*. In Algorithm 1, the set S_i is updated considering vehicles' travel times. First, a request vehicle (provider or consumer) that cannot satisfy the condition $\hat{t}_{p_v,i} < \gamma$ is removed from the set S_i. In this case, a provider cannot finish downloading the file using cellular links until it leaves the communication range of the OP, or a consumer will leave the OP before the file is ready. Second, a consumer in the set S_i is excluded when its expected content retrieval delay is longer than the tolerance time constraint δ. After pruning unsatisfiable vehicles, the DOVE algorithm selects an RN that covers the largest number of vehicles. Then, we select the first arriving vehicle at the OP as a provider from the set S_i^*. Finally, we can obtain the approximately minimum number of OPs with providers.

2) DOVE$^+$ Algorithm (using Multiple Donations): In the DOVE algorithm, each provider is scheduled to store the downloaded file only at a single OP for other consumers. However, providers can put the file at multiple OPs since they will pass multiple RNs. Thus, we can enhance the DOVE algorithm to take advantage of multiple donations to further reduce the number of providers that consume cellular traffic. Fig. 3 shows the basic idea using multiple donations (drawn by the gray lines). We call this enhanced algorithm **DOVE$^+$**. The following procedure of DOVE$^+$ algorithm is the same as the DOVE algorithm except step *(3) additional donations*:

(1) Select an S_i^* to decide an OP that directly covers

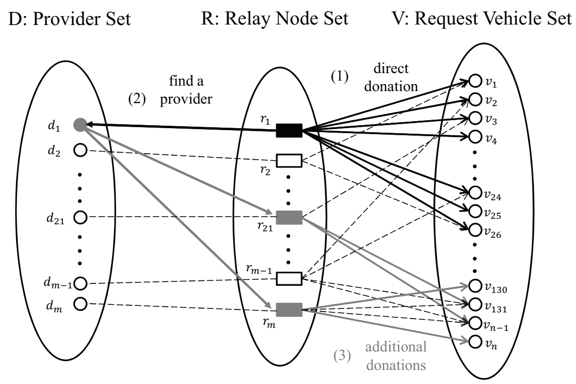

Fig. 3. The basic idea of multiple donations (DOVE$^+$).

the maximum vehicles (line 6 in Algorithm 1). In Fig. 3, RN r_1 is selected as an OP. Then, it can cover vehicles v_1-v_4 and v_{24}-v_{26}.

(2) Find a provider that is the first vehicle reaching the selected OP (line 7 in Algorithm 1). In Fig. 3, a vehicle d_1 is selected as a provider for RN r_1.

(3) Find additional consumers using RNs where a provider will pass (multiple donations). In Fig. 3, d_1 can store the file in RN r_{21} and r_m, so this can additionally cover vehicles v_{130}, v_{131}, v_{n-1}, and v_n.

(4) Repeat the steps (1)-(3).

Clearly, there is a trade-off between DOVE and DOVE$^+$ in terms of the reduced cellular traffic and the content retrieval delay. DOVE$^+$ algorithm reduces the aggregated usage of cellular links for vehicles while increasing the average content retrieval delay. We can select one depending on the requirements. Note that the time complexity of DOVE or DOVE$^+$ algorithms is computed as $O(VF \cdot min(V, F))$, which is a polynomial time [17].

D. The Selection of Providers

In this subsection, we describe how we select providers. When the requested file is located at the OP, the consumers passing through the OP can offload their cellular traffic to the OP by DSRC communications. To cover as many vehicles as possible, the TCC selects the first vehicle reaching the communication range of the OP as a provider. The arrival time of vehicles is calculated using the time prediction model, as described in Section III-A. Note that incentives for providers can be rewarded by mobile providers (e.g., reduced cost for using their cellular networks). The incentive policy for data offloading is left as future work.

E. The Operation of Vehicles using Offloading Positions

When a vehicle receives the offloading information from the TCC, it can know its OP and whether it is a provider or not. If a vehicle is selected as a provider, the vehicle downloads the file over cellular link. Then, the provider stores the file in the scheduled OP over DSRC link. On the other hand, if a vehicle is decided as a consumer, it postpones downloading the file. Then, the vehicle retrieves the file from the scheduled OP over DSRC link. Since the file is retrieved by DSRC link, the consumer can reduce the cost of cellular link usage. Note that vehicles should consume cellular traffic in the following cases: (i) vehicles are notified as providers by the TCC, (ii) vehicles fail to receive the file from the scheduled OP, and (iii) vehicles do not retrieve the file within the tolerance time.

IV. Performance Evaluation

In this section, we evaluate the performance of DOVE. Since we have no other state-of-the-art algorithm for OP selection, we compare the selection algorithm of DOVE (called DOVE) with the following: (i) a random selection algorithm that randomly selects OPs (called *Random*), and (ii) a greedy selection algorithm that selects an OP that covers the largest number of request vehicles without time consideration (called *Greedy*), and (iii) DOVE$^+$ algorithm (called *DOVE*$^+$).

TABLE I. Simulation Configuration

Parameter	Description
Road network	The number of intersections is 49. The area of the road map is 8.25km$\times$9km.
Communication range of DSRC	Communication range $R = 200$ meters. Bandwidth of the DSRC = 25 Mbps [1].
Number of vehicles (N)	The number of vehicles moving within the road network. The default N is 300.
Vehicle speed (v)	$v \sim N(\mu_v, \sigma_v)$ where $\mu_v = \{20, 25, ..., 60\}$ MPH and $\sigma_v = 5$ MPH [2], [16]. The maximum and minimum speeds of vehicles are $\mu_v + 3\sigma_v$ and $\mu_v - 3\sigma_v$, respectively. The default (μ_v, σ_v) is $(40, 5)$.
Deployment ratio (α)	The ratio α of the number of deployed RNs to the total number of intersections. The default α is 100%.
Tolerance time (δ)	The maximum (tolerable) delay of vehicles. The default δ is 600 sec (i.e., 10 min) [9].
Cellular downloading time (γ)	The content downloading time through a cellular network. Cellular downloading time γ is 47.9 sec. (Assumption: file size=12 MB, BW=2.1 Mbps)

To this end, we have developed a packet-level discrete event simulator using the SMPL [18]. Table I summarizes the detailed description of the simulation configuration. We use a road network which consists of 49 intersections. The layout of the road network is based on the map of Minneapolis downtown in Minnesota in the US. Each vehicle's movement pattern is determined by a Hybrid Mobility model of City Section Mobility model [19] and Manhattan Mobility model [20]. To reflect the stop sign or traffic signal, each vehicle waits for a random (uniform) waiting time between 0 and 10 sec at each intersection [2]. Each vehicle's speed is generated from a normal distribution of $N(\mu_v, \sigma_v)$ [2], [16], [21] as described in Table I. The communication-related parameters (e.g., DSRC communication range) in our evaluation are selected, based on a typical DSRC scenario [1]. To share a file in our simulations, each request vehicle sends a content request to the TCC along with its trajectory and current location information. Note that 5% of vehicles are selected as request vehicles that request the file. The total size of the request vehicle set is 1,000 and we use the averaged evaluation results. Unless otherwise specified, the default values of parameters in Table I are used.

A. Overall Performance of Data Offloading

We compare the offloading performance of DOVE, *Greedy*, *Random*, and *DOVE*$^+$. Note that we assume a perfect time-prediction solution of DOVE for comparison purposes. The

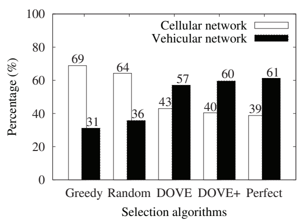

Fig. 4. The ratio of cellular network usage and vehicular network usage.

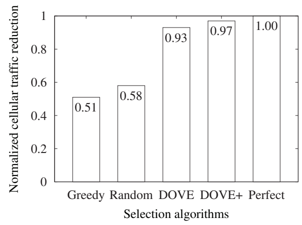

Fig. 5. The normalized cellular traffic reduction by offloading.

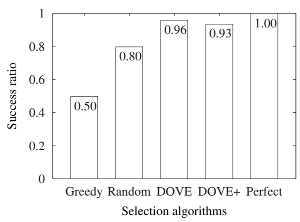

Fig. 6. The success ratio of offloading operations.

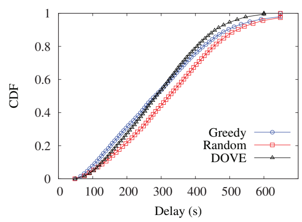

Fig. 7. The CDF of content retrieval delay of consumers.

perfect time-prediction solution (called *Perfect*) is assumed that the TCC can estimate the travel times of vehicles without any error caused by traffic signals. Thus, all the vehicles perfectly perform offloading operations according to the TCC's schedule.

As shown in Fig. 4, we analyze the portions of *the number of vehicles using the cellular network* and *the number of vehicles using the vehicular network* to the total number of request vehicles, respectively. When OPs are selected by DOVE, more than a half of request vehicles (i.e., 57%) offload their traffic from the cellular network to the vehicular network. Also, the offloading ratio of *DOVE*$^+$ is somewhat increased (i.e., 60%) since *DOVE*$^+$ reduces the number of providers by multiple donations. Compared to *Perfect* (61%), DOVE and *DOVE*$^+$ show almost comparable performance. However, only 31% and 36% of request vehicles take advantage of data offloading when they use *Greedy* and *Random*, respectively. These results indicate that DOVE selects more effective OPs compared to *Greedy* and *Random*. As a result, in Fig. 5, the reduced cellular traffic of DOVE (or *DOVE*$^+$) is close to that of *Perfect*. Note that *Greedy* selects an RN without time consideration. Thus, a lot of consumers fail to retrieve a file when providers cannot store it timely in the OPs. Similarly, *Random* does not consider the travel times of vehicles and randomly selects OPs so that it also shows the low offloading ratio. The result of *Greedy* reveals that a lot of consumers fail to receive a file, as providers cannot store it timely in the OPs. Thus, *Greedy* and *Random* show the reduction of about a half of cellular traffic, compared to *Perfect* as shown in Fig. 5.

Fig. 6 shows how many request vehicles successfully perform offloading operations according to the TCC's schedule. Vehicles fail in offloading operations when consumers cannot obtain the file from the scheduled OP within the tolerance time or providers cannot store the file in the OP. As shown in Fig. 6, DOVE and *DOVE*$^+$ show the high success ratio (96% and 93%, respectively) while *Greedy* and *Random* exhibit the lower success ratio. Because many request vehicles using *Greedy* or *Random* may fail to retrieve a file due to the above two types of the failures. In DOVE, about 4% of request vehicles fail in offloading operations due to the error of travel time prediction. This incorrect prediction is caused by a variation of waiting time due to traffic signals. Compared to DOVE, *DOVE*$^+$ shows a little lower success ratio (93%) since the prediction error of provider can cause more serious impact (i.e., higher failure ratio) on multiple donations. Note that about 50% of request vehicles using *Greedy* fail in offloading operations although *Greedy* chooses an RN that covers the largest number of request vehicles at each iteration. It shows the importance of travel time prediction. Thus, if TCC can perform better travel time prediction by real-time traffic measurement, our DOVE will provide a better data offloading.

Fig. 7 shows the cumulative distribution function (CDF) of content retrieval delays. We analyze the content retrieval delay of each consumer. Note that we omit the results of *DOVE*$^+$ and *Perfect* since they shows the similar performance to DOVE, and plotting their performance values hinders the readability of results. As shown in Fig. 7, all the consumers using DOVE obtain the file within the delay bound (i.e., 600 sec). However, *Greedy* and *Random* show the non-negligible portion of consumers that exceed the delay bound (3.2% and 4.3%, respectively). This is because *Greedy* and *Random* select OPs without considering the travel times of vehicles. Note that many vehicles using *Greedy* have the shorter content retrieval delay than DOVE. In DOVE, all vehicles are scheduled to utilize OPs considering the cellular downloading time γ (of a provider) to reduce the failure probability of offloading operations. However, *Greedy* selects OPs although travel times are less than the cellular downloading time, which leads to the fast content retrieval using cellular links due to the failure of offloading. As a result, about a half of consumers using *Greedy* show shorter content retrieval delay than vehicles using DOVE.

B. *The Impact of Vehicle Number*

Assume that the number of DOVE users is proportional to the total vehicle number. Then, the number of vehicles in the road network determines the number of vehicles participating in DOVE service. In this subsection, we show how many vehicles can offload their traffic from cellular links to the vehicular network, as the number of DOVE users increases. In our simulations, we assume that 5% of vehicles participate in DOVE service to retrieve the same content file. As more vehicles in the road network request the same popular file, both the number of overlapped RNs (candidates for OPs) and the number of consumers passing through each OP also increase together. Thus, a significant reduction in cellular traffic is expected when vehicles utilize the OPs for data offloading.

Fig. 8 shows the reduction of cellular traffic with varying the number of vehicles, that is, from 100 to 500. For comparison, we normalize the results since the number of vehicles requesting the same file is changed under different vehicle numbers. In Fig. 8, as the number of vehicles increases, DOVE and *DOVE*$^+$ reduce a significant amount of cellular traffic. However, *Greedy* and *Random* show the marginal difference in traffic reduction. This is because many vehicles using *Greedy* and *Random* fail in offloading operations, leading to the performance degradation. In contrast, most vehicles using DOVE and *DOVE*$^+$ successfully perform offloading operations since they select OPs based on the time prediction. These results indicate the growth in consumers passing through each

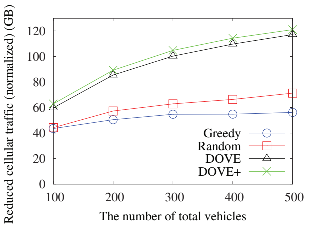

Fig. 8. The impact of vehicle number on the traffic reduction.

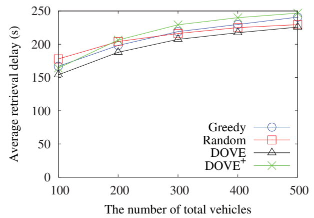

Fig. 9. The impact of vehicle number on the average retrieval delay.

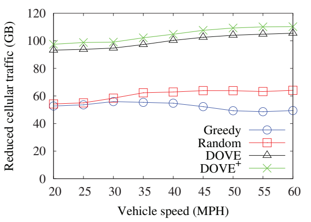

Fig. 10. The impact of vehicle speed on the traffic reduction.

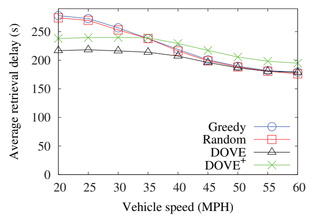

Fig. 11. The impact of vehicle speed on the average retrieval delay.

OP causes the performance degradation when providers fail to store the file at OPs.

To investigate the content retrieval delay of vehicles, we define the term *average retrieval delay* to be the average of the *content retrieval delays* of all vehicles. In Fig. 9, we notice that the average retrieval delays of all the algorithms increase as the number of vehicles increases. This is because many vehicles postpone content downloading and try to retrieve the file via vehicular networks, which increases the content retrieval delay of each consumer, leading to the longer average retrieval delay. However, the content retrieval delay of a vehicle in *Greedy* and *Random* can be longer than the expected travel time to the selected OP when vehicles fail in offloading operations and switch to cellular links. Also, the average retrieval delay of *DOVE*$^+$ is longer than DOVE since multiple donations of *DOVE*$^+$ increase vehicles' travel time to the scheduled OPs.

C. The Impact of Vehicle Speed

We investigate how the change of mean vehicle speed affects the performance of data offloading. Fig. 10 shows the reduced cellular traffic under different mean vehicle speeds. As shown in the Fig. 10, for DOVE, *DOVE*$^+$, and *Random*, the higher vehicle speed somewhat increases the reduced cellular traffic. This is because the high vehicle speed yields the shorter travel time of vehicles, so this increases the number of vehicles that can utilize the OPs within the delay bound (i.e., 600 sec). However, the high vehicle speed slightly decreases the reduced cellular traffic of *Greedy*. This is because *Greedy* selects an RN that covers the largest number of vehicles, so the failure of providers increases the failure of consumers exploiting OPs.

In Fig. 11, the higher vehicle speed decreases the average retrieval delay of all the algorithms. This is because the shorter travel time of vehicles decreases the average retrieval delay. Compared to DOVE, the lower vehicle speed highly increases the average retrieval delay of *Greedy* and *Random*. This is because a lot of vehicles cannot retrieve the file until the delay bound, leading to the increase of the content retrieval delay. Note that the average retrieval delay of *DOVE*$^+$ is longer than DOVE because of the longer travel time to the scheduled OPs, but the trends of two algorithms are the same.

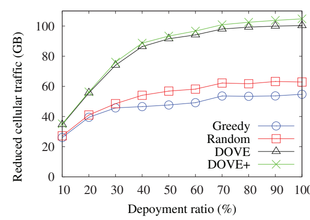

Fig. 12. The impact of partial deployment of RNs.

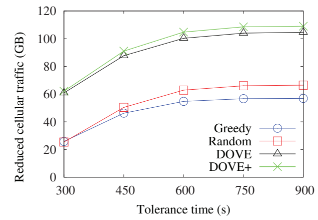

Fig. 13. The impact of tolerance time.

D. The Impact of Deployment Ratio and Tolerance Time

We investigate how the partial deployment of RNs and tolerance time affect the performance of DOVE. Fig. 12 shows the reduced cellular traffic under different deployment ratios, that is, from 10% to 100%. When the ratio is 100%, RNs are assumed to be deployed at all intersections. Note that we assume that RNs are deployed starting from the center toward the boundary of the road network. In all the algorithms, the higher deployment ratio leads to the less usage of cellular links. Compared to DOVE (or *DOVE*$^+$), *Greedy* and *Random* achieve lower performance gain since they cannot fully utilize the densely deployed RNs. Interestingly, DOVE shows that RNs in only 20% of intersections can reduce the 55.7 GB of cellular traffic, that is, about half of traffic reduction in the full deployment scenario (i.e., 100.3 GB). This result indicates our DOVE framework can achieve substantial data offloading even with the partial deployment scenario. In our simulation setting, the deployment ratio of 50% with DOVE algorithm shows the 10% of performance difference (91.5 GB) compared to the full deployment scenario.

Fig. 13 shows the reduced cellular traffic under different tolerance times, that is, from 300 sec to 900 sec. As shown in Fig. 13, DOVE and *DOVE*$^+$ show almost comparable performance, and the longer tolerance time increases the reduced cellular traffic in all the algorithms. This result implies that more tolerant users have more chances to offload their traffic from cellular links to RNs through DOVE. Also, longer tolerance time can increase the success probability of offloading operations since it can reduce the effect of prediction error caused by a variation of waiting time due to traffic signals. Therefore, it can be concluded that DOVE can provide cost-effective offloading service with infrastructure nodes for the driving safety.

V. Related Work

In vehicular networks, research about data forwarding and data dissemination has been investigated. *TBD* [22], *TSF* [2], *GeOpps* [23], and Leontiadis *et al.* [24] utilize vehicle trajectories along with vehicular traffic statistics to forward packets with shorter delivery delay and better delivery probability. For the data dissemination, Leontiadis *et al.* [25] propose a content based information dissemination protocol. Their protocol utilizes vehicle trajectories in order to disseminate some information to specific areas where vehicles need to receive it. Similarly, We *et al.* [26] propose *MDDV* that exploits a predefined trajectory for data dissemination. For all those existing approaches, authors focus on vehicle trajectories to forward/disseminate an information to vehicles in the specific location. On the other hand, DOVE investigates how to utilize both vehicle trajectories and potential benefits of the vehicular networks for offloading purposes.

There have been several studies to offload cellular traffic using WiFi or opportunistic communications [7]–[9], [27], [28]. Lee *et al.* [7] show that the offloading of 3G traffic to WiFi can significantly benefit mobile providers in terms of infrastructure cost. In *BreadCrumbs* [8], authors show that the forecast of WiFi access can improve the offloading of several applications. Also, recent studies [27]–[29] exploit mobile devices of social friends to offload cellular traffic through opportunistic communications. These previous studies focus on the data offloading using WiFi hotspots and client devices. In contrast, DOVE investigates the utilization of vehicular infrastructure (i.e, relay nodes) for data offloading.

Several works [6], [13], [30] address the mobile data offloading problem in vehicular networks. Li *et al.* [6] conduct the mathematical analysis with optimization problem of the offloading in vehicular networks. Also, Malandrino *et al.* [13] investigate a content prefetching at road-side units (RSUs) via the Internet. In [30], Siris and Kalyvas propose an offloading scheme that exploits WiFi hotpots located in single vehicle's route. For data offloading, above works focus on WiFi or RSUs, which are deployed with the Internet connectivity. Our work differs in that we utilize relay nodes, which do not have the Internet connectivity for cheaper installation cost, by considering vehicle trajectories of multiple vehicles to offload the redundant cellular traffic of vehicles.

VI. Conclusion

In this paper, we propose a Data Offloading framework through Vehicular nEtworks (DOVE). In DOVE, vehicle trajectories are utilized for data offloading. To select effective offloading positions (OPs), we formulate the selection of OPs as a spatio-temporal set-covering problem and then propose a time-prediction based set-covering algorithm. Simulation results show that DOVE can reduce about 57% of cellular link usage via OPs. We believe our DOVE will be used as one of solutions to resolve the mobile traffic explosion. As future work, we will investigate how to use RSUs as providers that can fetch the data from the Internet without using cellular links, and also how to deal with large-size content files for data offloading with multiple OPs.

Acknowledgment

This research was supported in part by Basic Science Research Program through the "National Research Foundation of Korea(NRF)" funded by the Ministry of Science, ICT & future Planning (2013R1A2A2A01016562) and in part by Global Research Laboratory Program (2013K1A1A2A02078326) through NRF. Note that Jaehoon (Paul) Jeong and Ted "Taekyoung" Kwon are the corresponding authors.

References

[1] Y. L. Morgan, "Notes on DSRC & WAVE Standards Suite: Its Architecture, Design, and Characteristics," *IEEE Communications Surveys & Tutorials*, vol. 12, no. 4, pp. 504–518, Oct. 2010.

[2] J. Jeong *et al.*, "TSF: Trajectory-based Statistical Forwarding for Infrastructure-to-Vehicle Data Delivery in Vehicular Networks," in *Proc. of IEEE ICDCS*, 2010.

[3] J. Chi *et al.*, "An Effective RSU Allocation Strategy for Maximizing Vehicular Network Connectivity," *IJCA*, vol. 6, no. 4, pp. 259–270, Aug. 2013.

[4] Paul Stenquist, "With Its Eye on Infotainment, General Motors Is Upgrading OnStar," http://wheels.blogs.nytimes.com/2013/02/25/with-its-eye-on-infotainment-general-motors-is-upgrading-onstar/.

[5] "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2013-2018," Feb. 2014.

[6] Y. Li *et al.*, "Coding or Not: Optimal Mobile Data Offloading in Opportunistic Vehicular Networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 1, pp. 318–333, 2014.

[7] K. Lee *et al.*, "Mobile data offloading: How much can WiFi deliver?" in *Proc. of ACM CoNEXT*, 2010.

[8] A. J. Nicholson and B. D. Noble, "BreadCrumbs: Forecasting mobile connectivity," in *Proc. of ACM MOBICOM*, 2008.

[9] Y. Im *et al.*, "Amuse: Empowering users for cost-aware offloading with throughput-delay tradeoffs," in *Proc. of IEEE INFOCOM*, 2013.

[10] S. Woo *et al.*, "Comparison of caching strategies in modern cellular backhaul networks," in *Proc. of ACM MobiSys*, 2013.

[11] PATH Project, "Investigating the Potential Benefits of Broadcasted Signal Phase and Timing (SPaT) Data under IntelliDrive," University of California, Berkeley, Tech. Rep., May 2011.

[12] Philadelphia Department of Transportation, "Traffic Control Center," http://philadelphia.pahighways.com/philadelphiatcc.html.

[13] F. Malandrino *et al.*, "Offloading cellular networks through its content download," in *Proc. of IEEE SECON*, 2012.

[14] Jupiter Research, "Municipal Wireless: Partner to Spread Risks and Costs While Maximizing Benefit Opportunities," Jupiter Research, Tech. Rep., Jun. 2005.

[15] Garmin Ltd., "Garmin Traffic," http://www8.garmin.com/traffic/.

[16] D. S. Berry and D. M. Belmont, "Distribution of Vehicle Speeds and Travel Times," in *Proc. of the Second Berkeley Symposium on Mathematical Statistics and Probability*, Jul. 1950.

[17] U. Feige, "A threshold of ln n for approximating set cover," *J. ACM*, vol. 45, no. 4, pp. 634–652, Jul. 1998.

[18] M. MacDougall, *Simulating Computer Systems: Techniques and Tools*. MIT Press, 1987.

[19] T. Camp *et al.*, "A Survey of Mobility Models for Ad Hoc Network Research," *WCMC: Special Issue on Mobile Ad Hoc Networking: Research, Trends and Applications*, vol. 2, pp. 483–502, Aug. 2002.

[20] F. Bai *et al.*, "IMPORTANT: A framework to systematically analyze the Impact of Mobility on Performance of RouTing protocols for Adhoc NeTworks," in *Proc. of IEEE INFOCOM*, 2003.

[21] V. Muchuruza and R. Mussa, "Traffic Operation and Safety Analyses of Minimum Speed Limits on Florida Rural Interstate Highways," in *Proc. of MIDCON*, 2005.

[22] J. Jeong *et al.*, "Trajectory-Based Data Forwarding for Light-Traffic Vehicular Ad-Hoc Networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 22, no. 5, pp. 743–757, May 2011.

[23] I. Leontiadis and C. Mascolo, "GeOpps: Geographical Opportunistic Routing for Vehicular Networks," in *Proc. of IEEE WoWMoM*, 2007.

[24] I. Leontiadis *et al.*, "Extending Access Point Connectivity through Opportunistic Routing in Vehicular Networks," in *Proc. of IEEE INFOCOM*, 2010.

[25] I. Leontiadis, P. Costa, and C. Mascolo, "A Hybrid Approach for Content-based Publish/Subscribe in Vehicular Networks," *Elsevier Pervasive and Mobile Computing*, vol. 5, no. 6, pp. 697–713, Dec. 2009.

[26] H. Wu *et al.*, "MDDV: A Mobility-Centric Data Dissemination Algorithm for Vehicular Networks," in *Proc. of ACM VANET*, 2004.

[27] B. Han *et al.*, "Cellular traffic offloading through opportunistic communications: a case study," in *Proc. of ACM CHANTS*, 2010.

[28] Y. Li *et al.*, "Multiple mobile data offloading through delay tolerant networks," in *Proc. of ACM CHANTS*, 2011.

[29] X. Wang *et al.*, "Content dissemination by pushing and sharing in mobile cellular networks: An analytical study," in *Proc. of IEEE MASS*, 2012.

[30] V. A. Siris and D. Kalyvas, "Enhancing mobile data offloading with mobility prediction and prefetching," in *Proc. of ACM MobiArch workshop*, 2012.