Poster: Greedy Data Dissemination Algorithm for Infrastructure-to-Vehicle Services

Inshick Kim
Vehicle Infotainment System Lab.
LG Electronics
Seoul, Republic of Korea
Email: inshick.kim@lge.com

Ryangsoo Kim, Hyuk Lim*
Gwangju Institute of Science and Technology (GIST)
Gwangju, Republic of Korea
Email: {rskim, hlim}@gist.ac.kr

Kyung-Joon Park
Daegu Gyeongbuk Institute of Science and Technology (DGIST)
Daegu, Republic of Korea
Email: kjp@dgist.ac.kr

***Abstract*—This paper considers a data dissemination problem for vehicular cloud systems, in which delivery services from a data center provide data to vehicles through roadside wireless access points having local data storages. We propose a greedy data prefetching algorithm; by exploiting the completely predictable route information and the stochastic characteristics of the communication with APs, the algorithm predetermines how to distribute a set of data from a data center to local data storages, and minimizes the amount of data dissemination. The results from extensive simulations show that the proposed scheme can achieve efficient data dissemination in a variety of vehicular scenarios while minimizing the data transfer.**

***Index Terms*—Vehicular networks, data dissemination, data prefetching.**

I. Introduction

We consider an infrastructure-to-vehicle (I2V) communication for providing cloud services such as those providing traffic information and file downloading. The vehicular network infrastructure for vehicular cloud services consists of cloud servers at a data center and roadside wireless APs with small memory buffers. The data center is connected to roadside wireless APs having local data storages, and mobile vehicles can access the data center through a nearby wireless AP. However, some vehicles may not have continuous access to the data center, especially when APs are sparsely distributed in the area of interest and/or the transmission coverage between nearby APs may not overlap. Thus, it is necessary to expedite data dissemination from the data center to the roadside wireless APs having local data storages prior to APs being requested from vehicular subscribers and to divide the data into a number of small chunks—which are the basic unit for reliable data delivery. Recently, online navigation based on GPS has been widely used, and thus drivers can be easily aware of their direct route to a specific destination. The vehicle routes traced in the online navigation device are delivered to the data center through the vehicular cloud system. Then, the data center utilizes the vehicle routes to expedite the data dissemination to vehicles.

To date, there have been a number of attempts made to address data dissemination issues in vehicular networks. For example, Hao *et al.* [1] proposed an MDDV algorithm based on data being geographically forwarded along a predefined trajectory in a partitioned and highly mobile vehicular network, and Zhao *et al.* [2] studied data dissemination between vehicles and roadside infrastructure, in a system that stored and rebroadcast buffered data. Lee *et al.* [3] described MobEyes, which is a vehicular sensing architecture for sensed urban data such as video clip dissemination and harvesting in vehicular sensor networks.

In a data dissemination problem, in which delivery services from a data center provide data to vehicles through roadside wireless access points having local data storages, if all local data storages have the entire data chunks for every vehicle, then vehicles are able to successfully receive their requested data. However, though successful, this method may not be efficient because all date chunks need to be copied to every local data storage. Moreover, the data storage size of these local replicas are limited and it takes a long time to disseminate all the chunks in a data center, and if there are no requests from vehicular subscribers the replicas do not need to be copied from the data center.

II. Proposed Data Dissemination Scheme

We propose a vehicle route based greedy data prefetching algorithm. By exploiting the completely predictable route information and the stochastic characteristics of the communication with APs, our proposed algorithm decides how to distribute in advance the data of interest from a data center to local data storages using a minimum transfer of data.

For efficient cloud resource usage of storage space in local data storages and network bandwidth in the data center, it is desirable to minimize the amount of data transferred from the data center to wireless APs, as long as a certain level of of data dissemination quality is preserved. Let w and u denote the numbers of wireless APs and data chunks, respectively. Also, let $\mathbf{X} = (x_{a,b})_{w\times u}$ denote a binary decision matrix where $x_{a,b}$ = 1 if the b-th chunk is to be prefetched to the a-th wireless AP, and otherwise is 0. Because the data chunk transferred to the data storages corresponds to the sum of chunk occupancy at every local data storage, it can be formulated as follows:

$$\begin{aligned} \text{minimize} \quad & \mathbb{1}_w^T \cdot \mathbf{X} \cdot \mathbb{1}_u \\ \text{subject to} \quad & \|\mathbf{R} \odot \mathbf{P}\|_{max} \le \epsilon \\ & \mathbf{X} \in \{0,1\}^{w\times u}, \end{aligned} \qquad (1)$$

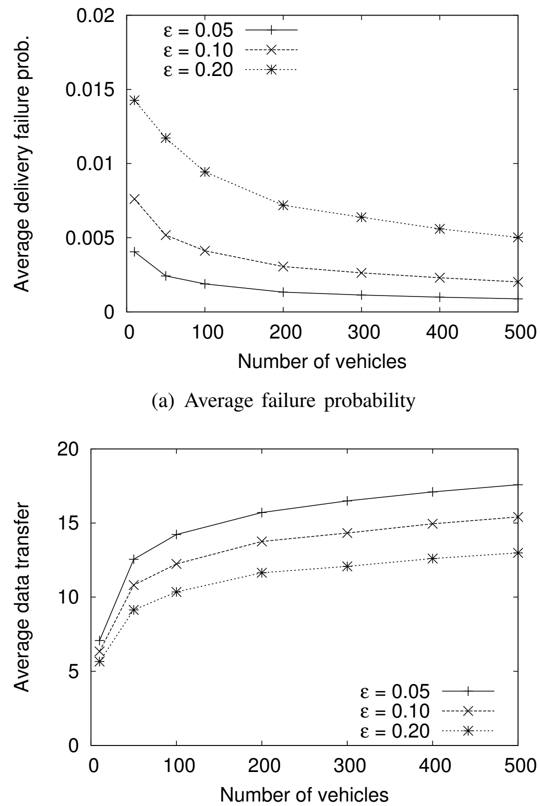
Algorithm 1 Proposed greedy algorithm

// Index set $\mathcal{A}_{\mathbf{x}_j}$: $\mathcal{A}_{\mathbf{x}_j} = \{i : x_{i,j} = 0, x_{i,j} \in \mathbf{x}_j, \forall i\}$
// Objective function $f(\mathcal{A}_{\mathbf{x}_j})$: $f(\mathcal{A}_{\mathbf{x}_j}) = w - \mathbb{1}_w^T \cdot \mathbf{x}_j$
for j= 1 to u **do**
  // Initialization
  $\mathbf{x}_j = \arg\min_{\mathbf{x}_j \in \{0,1\}^{w\times 1}, |\mathcal{A}_{\mathbf{x}_j}|=3} \|\mathbf{r}_j \odot \mathbf{p}_j(\mathcal{A}_{\mathbf{x}_j})\|_{max}$
  // Main loop
  while $\mathcal{A}_{\mathbf{x}_j} \neq \mathcal{A}_{\mathbf{0}_{w\times 1}}$ **do**
    Find the index i^*
      $= \arg\min_{i^* \in \mathcal{A}_{w\times 1} \setminus \mathcal{A}_{\mathbf{x}_j}} \|\mathbf{r}_j \odot \mathbf{p}_j(\mathcal{A}_{\mathbf{x}_j} \cup \{i^*\})\|_{max}$
    if $\|\mathbf{r}_j \odot \mathbf{p}_j(\mathcal{A}_{\mathbf{x}_j} \cup \{i^*\})\|_{max} \leq \epsilon$ **then**
      Stop the while loop
    end if
    $\mathcal{A}_{\mathbf{x}_j} \leftarrow \mathcal{A}_{\mathbf{x}_j} \cup \{i^*\}$
    $\mathbf{x}_j[i^*] \leftarrow 0$
  end while
end for

where $\mathbb{1}_l \in \mathbb{R}^l$ is an all-ones vector, $\mathbf{R}$ is a chunk request matrix, $\mathbf{P}$ is a dissemination failure probability matrix, ϵ is a tolerable dissemination failure probability for each pair of vehicles and their requested chunks, and $\odot$ is Hadamard product operator. Note that $\mathbf{P} = (p_{a,b})_{v\times u}$, where $p_{a,b}$ is the probability that the a-th vehicle fails to download the b-th chunk because none of wireless APs on the a-th vehicle's route has the b-th chunk. A constraint is then imposed to retain a minimum dissemination failure probability. The optimal solution of $\mathbf{X}$ in (1) minimizes the total amount of data transferred to the APs while keeping the dissemination failure probabilities for all pairs of vehicles and their requested data chunks below ϵ. In order to apply this optimization to large vehicular systems at a lower computational complexity, a greedy algorithm is proposed as shown in Algorithm 1. The total number of iterations of the greedy algorithm is bounded by $(w^2 \times u)$.

III. Simulation Results

In order to evaluate the performance of the proposed greedy algorithm, we conduct simulations using MATLAB. In brief, we implement the data prefetching scheme with regard to increases in the number of vehicles, wireless APs, and chunks under several values of ϵ. Figure 1(a) shows the average failure probability of vehicle-chunk pairs for three different ϵ. In the figure, when ϵ is 20%, the average failure probability is the highest, thereby implying that if the tolerable dissemination failure probability is high, the failure probability is also likely to be higher; when the number of vehicles is increased, the average failure probability also decreases. Specifically, when ϵ is small and the number of vehicles is large, the graph of the average failure probability clearly decreases. Figure 1(b) shows the amount of average data transferred of APs using the three different ϵ. In this depiction, when ϵ is small, the graph of the amount of data transfer is high, as opposed to the case of Figure 1(a), as the amount of data transfer affects the failure probability. In other words, if the amount of the data transfer is high, then it can reduce the average failure probability, to some degree.



(a) Average failure probability

(b) Data transfer

Fig. 1. Simulation results for our greedy algorithm with respect to the number of vehicles when w = 50 APs, u = 100 chunks, and ϵ = 5%, 10%, 20%.

IV. Conclusion

In this paper, we proposed a greedy algorithm to minimize the amount of data transferred from a data center to the wireless APs. Through the simulations, we verified that the proposed scheme could achieve efficient data dissemination in a variety of vehicular scenarios, by minimizing the data transfer. In the future, we plan to extend our data prefetching algorithm to more specific and realistic applications.

Acknowledgment

This work was supported in part by the NRF funded by MSIP of the Korea government (2014R1A2A2A01006002), and by the ICT R&D program of MSIP and IITP of the Korea government (B0101-15-0557, Resilient CPS Research).

References

[1] H. Wu, R. Fujimoto, R. Guensler, and M. Hunter, "MDDV: A mobility-centric data dissemination algorithm for vehicular networks," in *Proc. of ACM International Workshop on Vehicular Ad-hoc Networks*, 2004.

[2] J. Zhao, Y. Zhang, and G. Cao, "Data pouring and buffering on the road: A new data dissemination paradigm for vehicular ad hoc networks," *IEEE Trans. on Vehicular Technology*, vol. 56, no. 6, pp. 3266–3277, 2007.

[3] U. Lee, E. Magistretti, M. Gerla, P. Bellavista, and A. Corradi, "Dissemination and harvesting of urban data using vehicular sensing platforms," *IEEE Trans. on Vehicular Technology*, vol. 58, no. 2, pp. 882–901, 2009.