Relay Selection Strategy (RSS) Design for In-Vehicle Storage (IVS) System

Binbin Hu*†, Rongqing Zhang‡, Luoyang Fang§, Xiang Cheng*†, and Liuqing Yang§

* State Key Laboratory of Advanced Optical Communication Systems & Networks,
School of EECS, Peking University, Beijing 100871, China

† Key Laboratory of Wireless Sensor Network & Communication, Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, 865 Changning Road, Shanghai 200050, China

‡ School of Software Engineering, Tongji University, Shanghai 201804, China

§ Department of Electrical and Computer Engineering, Colorado State University, Fort Collins, CO 80523, USA

***Abstract*—In recent years, autonomous driving has attracted a vast amount of attention from both industry and academia, which has introduced large amounts of region-related data. In order to release the burden in core networks caused by the communication demands for such data, various in-vehicle storage (IVS) systems have been widely studied to bring contents closer to users. To cope with the mobility issue hindering the realization of IVS systems, in this paper, we propose a relay selection strategy (RSS) consisting of the relay map construction (RMC) algorithm and the relay pair matching (RPM) algorithm with the assistance of the vehicle route information. By considering both the potential transmission amount and the waiting time, the proposed RSS generates an overall optimal relay assignment for the IVS system. The performance gain of the proposed RSS compared with the baseline is evaluated by a realistic simulator in terms of the relay failure ratio, the retrieval throughput, and the RSU consumption ratio. Simulation results show that the proposed RSS achieves higher efficiency and robustness than the baseline.**

I. Introduction

Autonomous driving has recently attracted a vast amount of attention from both industry and academia. In order to achieve effectiveness and robustness, recent advancement of data processing in this field has introduced large amounts of region-related data with the assistance of publicly available datasets, e.g., *Cityscapes* [1], *KITTI Vision Benchmark Suite* [2], and *Daimler Urban Segmentation* [3]. One of the most common forms of region-related data is widely used high-definition (HD) maps [4] consisting of 3D elevation model [5], detected lane information [6], and stationary map [7]. The communication demands from such region-related data in vehicular networks result in a burden to core networks, which in turn leads to quality of experience (QoE) degradation of vehicular users. In cellular networks, the potential of in-device storage is explored to alleviate the burden on core networks by bringing contents closer to users. In these works, storage nodes are always assumed to be static or low-mobility, modeled by various Poisson process models [8] [9], which is clearly insufficient to deal with practical high mobility in vehicular networks. In order to alleviate the backhaul link congestion in vehicular networks, most literature focused on storing contents of interest at road side units (RSUs) [10]–[12] without the network throughout gain by spatial frequency reuse.

This work was in part supported by the Ministry National Key Research and Development Project under Grant 2017YFE0121400, the National Science and Technology Major Project under Grant 2018ZX03001031, the Major Project from Beijing Municipal Science and Technology Commission under Grant Z181100003218007, the open research fund of Key Laboratory of Wireless Sensor Network & Communication under Grant 2017003, Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, the National Natural Science Foundation of China under Grants 61622101 and 61571020, and the National Science Foundation under Grant CPS-1932413.

A few existing schemes also studied in-vehicle caching [13] [14] but with insufficient and impractical considerations on vehicle mobility, which is inadequate for region-related data storage. The realization of in-vehicle storage has mainly suffered from vehicle mobility, as a vehicle with stored data can only physically reside in a region of interest for a short period of time, except for the situations of heavy traffic jams or congestions. In our previous work [15]–[17], we proposed a distributed in-vehicle storage system based on data relays via vehicle-to-vehicle (V2V) communications between leaving vehicles and incoming vehicles with minor assistance from RSUs. However, with simplified consideration of road conditions, this straightforward data relay approach may lead to challenging implementation in realistic situations. The reduction of data relay efficiency is mainly due to the complicated road conditions and unpredictable vehicle movements, which makes it making it arduous to efficiently assign an appropriate vehicle for the leaving vehicle to relay its storage. Therefore, we intend to employ the predicted vehicle route information [18]–[20] to enhance the relay selection design.

In this paper, with the assistance of introduced vehicle route information, we propose a novel relay selection strategy (RSS) to improve the efficiency of the data relay in the in-vehicle storage (IVS) system. Specifically, the RSS consists of two core algorithms, namely the relay map construction (RMC) algorithm and the relay pair matching (RPM) algorithm, which are designed to evaluate the relay success probability and to maximize the overall relay success, respectively. The main function of the RMC algorithm is to construct a relay map indicating the relay success probability between vehicles, by calculating the potential transmittable data amount based on the predicted vehicle route information. With the constructed relay map including both the potential transmission amount and the waiting time, we formulate the relay selection problem as a bigraph maximum matching problem, and further provide the Hungarian method-based RPM algorithm to efficiently obtain an optimal relay assignment solution. The performance gain of our proposed RSS compared with the previous straightforward relay approach is verified by a simulator consisting of realistic road network and vehicular traffic, in terms of the relay failure ratio, the retrieval throughput, and the RSU consumption ratio.

The rest of this paper is organized as follows. In Section

978-1-7281-0962-6/19/$31.00 ©2019 IEEE

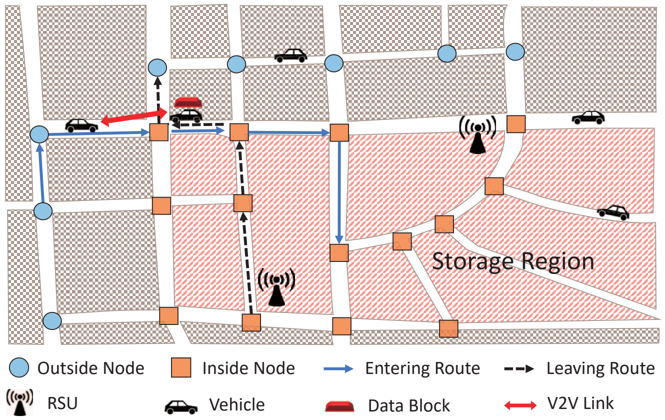

Fig. 1. Scenario for the IVS system in a multi-intersection multi-lane region.

II, we present our system model and briefly describe the relay selection problem. In Section III, we detail our proposed RSS consisting of the RMC algorithm and the RPM algorithm. We demonstrate the effectiveness and robustness of our proposed RSS through simulations in Section IV. Finally, Section V concludes the paper.

II. System Model and Problem Formulation

In this section, we first present the architecture as well as the system model of the investigated IVS system in vehicular networks. Then, we give an overview of the relay selection problem that we endeavor to solve in this paper.

A. Scenario Description

We consider a multi-tier hierarchical vehicular network in a multi-intersection multi-lane region without loss of generality, as shown in Fig. 1. Let $\mathcal{N} = \left\{n_1, n_2, \cdots n_{|\mathcal{N}|}\right\}$ and $\mathbf{G}_{|\mathcal{N}|\times|\mathcal{N}|}$ denote the set of road nodes representing road intersections, and the $|\mathcal{N}| \times |\mathcal{N}|$ matrix represents the road segments connecting the road nodes. The road nodes $\mathcal{N}$ can be divided into two disjoint sets, $\mathcal{N}_{in}$ and $\mathcal{N}_{out}$, representing the nodes inside the storage region and the nodes outside the storage region, respectively, that is, $\mathcal{N}_{in} \cup \mathcal{N}_{out} = \mathcal{N}$ and $\mathcal{N}_{in} \cap \mathcal{N}_{out} = \emptyset$. The element $g_{i,j}$ in $\mathbf{G}$ indicates the length of the road segment connecting node n_i and node n_j iff (if and only if) they are adjacent in the road map. A vehicle V **inside** the storage region $\mathcal{S}$ is defined as that it is in the road segment $g_{i,j}$ with both nodes inside the storage region, i.e.,

$$V \text{ in } \mathcal{S} \triangleq \{V \text{ in } g_{i,j} | n_i \in \mathcal{N}_{in}, n_j \in \mathcal{N}_{in}\}. \tag{1}$$

A vehicle V **outside** the storage region $\mathcal{S}$ is defined as that it is in the road segment $g_{i,j}$ with at least one node outside the storage region, i.e.,

$$V \text{ out } \mathcal{S} \triangleq \{V \text{ in } g_{i,j} | n_i \in \mathcal{N}_{out} \text{ or } n_j \in \mathcal{N}_{out}\}. \tag{2}$$

Based on the above representations, substantially all road conditions can be represented as maps consisting of road nodes and road segments, which eliminates the constraint of the road conditions in our previous work [15]–[17].

B. IVS System Framework

The IVS system is covered by RSUs, as shown in Fig. 1. Two types of wireless communication links, V2R links and V2V links with effective communication range R_{max}, are assumed in this paper. In the IVS system, the region-related data is stored in the vehicles inside the storage region. Therefore, vehicles in the storage region can retrieve region-related data from their own storage or via V2V links with vehicles around them. In order to maintain the survival of the region-related data in the storage region, the data distributing mechanism consists of two components as follows.

1) Region-related Data Processing: In order to adapt to the IVS system, in the backend server, the region-related data is first processed into an appropriate amount of data blocks by a method similar to the one in [17]. The difference is that, with the adaptability introduced by the proposed RSS, the data blocks are no longer restricted to be equal-sized. Specifically, we assume that the region-related data library consists of N_f files with request probability distribution modeled as Zipf distribution [10] as follows,

$$f_i(\gamma_r) = \frac{1/i^{\gamma_r}}{\sum_{j=1}^{N_f} 1/j^{\gamma_r}}, \ i \in \{1, 2, \cdots, N_f\}, \tag{3}$$

where $f_i(\gamma_r)$ denotes the request probability of file F_i, and the exponent γ_r is the parameter characterizing users' requests for the library. To assign higher availability to the file with higher popularity, all files are first divided into data chunks, which are then encoded by an MDS code with different coding redundancy. The redundancy level Ω_i of file F_i is dependent on the popularity rank i and tuned by a controllable parameter γ_c based on the Zipf distribution, i.e.,

$$\Omega_i \propto \frac{1/i^{\gamma_c}}{\sum_{j=1}^{N_f} 1/j^{\gamma_c}}. \tag{4}$$

In order to allocate the processed data to the vehicles in the storage region, the coded data chunks are combined as a set of data blocks $\mathcal{B} = \left\{B_1, B_2, \cdots B_{|\mathcal{B}|}\right\}$, where B_i denotes the size of the i-th data block. Each data block is then allocated to a different vehicle in the storage region via V2R link.

2) Data Block Relay: Since the data blocks will be lost with the vehicles leaving the storage region, in order to retain the lost data blocks, the core mechanism of the IVS system is the relay process. In the relay process, the vehicles leaving the storage region are supposed to relay their data blocks to the incoming vehicles. To provide a space for the data relay, in our previous work [15]–[17], we designed two transfer regions placed at the entrances/exits of the storage region. However, this approach requires a lot of manual work to design transfer regions, which is inefficient and hard to implement in the complicated road conditions. For the design of more efficient and general relay selection strategy, in this paper, we employ two waiting queues for the leaving vehicles and the incoming vehicles, and introduce the predicted vehicle route information to cope with complicated road conditions. The leaving vehicles and the incoming vehicles will be added to the leaving vehicle

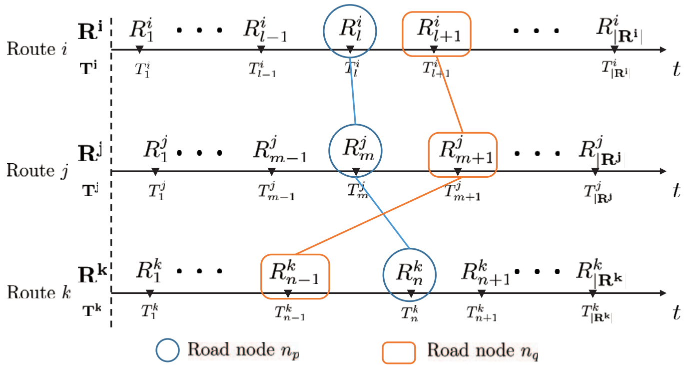

Fig. 2. An example of vehicles meeting in the same direction and in different directions, i.e., route i, j and route j, k.

quere $\mathbf{Q^{out}} = \left[\begin{array}{ccc} q_1^{out} & \cdots & q_{|\mathbf{Q^{out}}|}^{out}\end{array}\right]^T$ and the incoming vehicle quere $\mathbf{Q^{in}} = \left[\begin{array}{ccc} q_1^{in} & \cdots & q_{|\mathbf{Q^{in}}|}^{in}\end{array}\right]^T$, respectively, i.e.,

$$V \rightarrow \mathbf{Q^{out}} \triangleq \{V \text{ out } \mathcal{S} : (t + \Delta t) | V \text{ in } \mathcal{S} : t\}, \tag{5a}$$

$$V \rightarrow \mathbf{Q^{in}} \triangleq \{V \text{ in } \mathcal{S} : (t + \Delta t) | V \text{ out } \mathcal{S} : t\}. \tag{5b}$$

Each element $q_i^{out/in}$ of both queues denotes a vehicle member consisting of the waiting time $w_i^{out/in}$ from added to the queue and the route information. The route information denoted as $\langle \mathbf{R} \quad \mathbf{T} \rangle$ consists of location information $\mathbf{R}$ and corresponding time information $\mathbf{T}$, as shown in Fig. 2. For the location information, $\mathbf{R} = \left[\begin{array}{ccc} R_1 & \cdots & R_{|\mathbf{R}|}\end{array}\right]^T$ is a $|\mathbf{R}| \times 1$ vector, where R_i denotes the road node that the vehicle passes, i.e., $R_i \in \mathcal{N}$. The time information, $\mathbf{T} = \left[\begin{array}{ccc} T_1 & \cdots & T_{|\mathbf{R}|}\end{array}\right]^T$ is also a $|\mathbf{R}| \times 1$ vector, where T_i denotes the corresponding time point when the vehicle passes the road node R_i.

The proposed RSS matches the leaving vehicles in $\mathbf{Q^{out}}$ and the incoming vehicles in $\mathbf{Q^{in}}$ based on the relay weight, which will be presented in detail in Section III. The matched vehicles are assigned for the relay process and then will be removed from the corresponding waiting queue. To limit the waiting time in the queue, the maximum waiting time is set to w_{max}, which means that a vehicle will be removed from the queue if its waiting time $w_i^{out/in}$ exceeds w_{max}. Therefore, there are two conditions that can lead to the relay failure. First, as the V2V links may be opportunistic and not reliable, the relay process may be uncompleted. Second, a vehicle may be unmatched due to poor relay condition till getting removed from the waiting queue, which causes the region-related data loss. To remedy the data loss, once the relay process fails, the backend server will repair the lost data by resending a new data block to the storage region via V2R links.

C. Problem Description

In order to decrease the cost caused by relay failure, the problem of minimizing the overall relay failure can be formulated by considering the selection of the target relay vehicles. Let $\mathbf{M}$ denote the matching assignment, that is, the leaving vehicle q_i^{out} is assigned to relay its data to the incoming vehicle $q_{\mathbf{M}(i)}^{in}$. The object function of our optimization problem can be formulated as follows

$$\min_{\mathbf{M}} \quad \sum_{i=1}^{|\mathbf{Q^{out}}|} \mathcal{H}\left\{B_i - \sum_{k=1}^{K} \Delta t_k \cdot c_k\left(i, \mathbf{M}(i)\right)\right\} \tag{6a}$$

$$\text{s.t.} \quad \mathbf{M}(i) \neq \mathbf{M}(j), \forall\, i \neq j, \tag{6b}$$

$$\mathbf{M}(i) \leq \left|\mathbf{Q^{in}}\right|. \tag{6c}$$

where B_i denotes the data block size of leaving vehicle q_i^{out}, $c_k\left(i, \mathbf{M}(i)\right)$ denotes the transmission rate between vehicle q_i^{out} and vehicle $q_{\mathbf{M}(i)}^{in}$, and $\mathcal{H}(\cdot)$ is the Heaviside step function given by

$$\mathcal{H}(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \tag{7}$$

The term inside the Heaviside step function in Equation (6a) indicates the difference between the data block size and the transmission amount. The constraint (6b) limits that a leaving vehicle can only choose one incoming vehicle to relay data, which ensures no collision of duplicate data relays.

Algorithm 1: Relay Map Construction

Input: $\mathbf{Q^{out}}$ and $\mathbf{Q^{in}}$
Output: Relay map $\mathbf{A} = \left(\mathbf{Q^{out}}, \mathbf{Q^{in}}, \mathbf{E}\right)$

```
initialization;
for each q_i^out in Q^out and q_j^in in Q^in do
    for each R_m^i in R^i and R_n^j in R^j do
        if R_m^i = R_n^j then
            switch R_{m+1}^i do
                case R_{n-1}^j \\ diff dir. do
                    calculate Ê in equation (9) ;
                end
                case R_{n+1}^j \\ same dir. do
                    calculate Ê in equation (12) ;
                end
            end
            ξ ← α w_i^out / w_max + β Ê / B_i ;
            if ξ > 1 then
                e_{i,j} ← 1;
                break;
            end
        end
    end
end
```

III. Relay Selection Strategy Design

Targeting on the optimization problem formulated in (6), the RSS is carefully designed with the assistance of vehicle route information. The RSS consists of the RMC algorithm and the RPM algorithm aiming to estimate the relay success probability and to maximize the number of successful relays, respectively.

A. Relay Map Construction

According to Equation (6a), the term $\sum_{k=1}^{K} \Delta t_k \cdot c_k\left(i, \mathbf{M}(i)\right)$ represents the transmission amount between the leaving vehicle

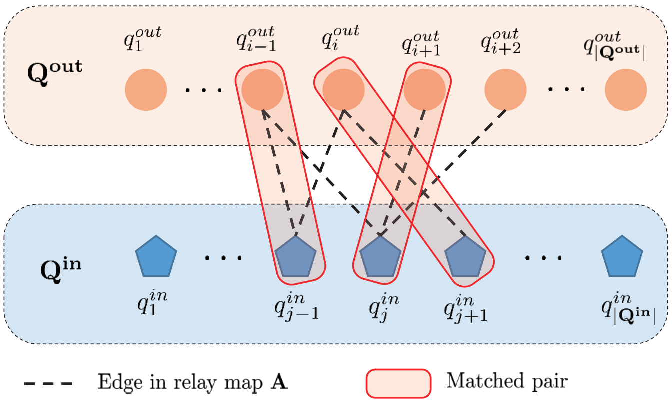

Fig. 3. An example of relay map and matched pairs.

q_i^{out} and the incoming vehicle $q^{in}_{\mathbf{M}(i)}$, which is unknown untill the end of the transmission. Therefore, the potential transmission amount between leaving vehicles and incoming vehicles should be estimated before making the relay assignment. In order to calculate the potential transmission amount $\hat{E}$ between any two vehicles, the vehicle route information is essential, since the relative distance between vehicles should be predicted first.

The detailed RMC algorithm is shown in Algorithm 1. In the process of calculating the potential transmission amount $\hat{E}$, all the cases of relative movements between vehicles are considered based on the vehicle route information. The first case is that vehicles meet in different directions, as the route j and route k in Fig. 2 show, i.e., $R^j_m = R^k_n$ and $R^j_{m+1} = R^k_{n-1}$. Both route j and route k contain the same consecutive road node n_p and n_q, which means that the two vehicles must pass the same road segment. However, route j enters the road segment from n_p while route k enters the road segment from n_q. Therefore, the two vehicles drive in the same segment in different directions. In this situation, we can obtain their meeting time as follows,

$$\begin{aligned} t^{\text{diff}} &= \left|T^j_m - T^j_{m+1}\right| + \left|T^k_n - T^k_{n-1}\right| \\ &- \left(\max\left(T^j_{m+1}, T^k_n\right) + \min\left(T^j_m, T^k_{n-1}\right)\right). \end{aligned} \tag{8}$$

According to the V2V communication model in [15], the potential transmission amount can be calculated considering the average transmission rate as follows,

$$\hat{E} = t^{\text{diff}} \frac{W\log_2(1+\eta D^{-4}) + W\log_2(1+\eta D^{-4}_{\max})}{2}, \tag{9}$$

where W is the channel bandwidth, η is the signal-to-noise rate (SNR) at transmitters, D is the lane width, and $D_{\max}$ is the maximum distance between two vehicles during the communication process. Note that $D_{\max}$ may not equal to the maximum V2V communication range $R_{\max}$ due to the short meeting time. Thus, assuming the uniform motion of vehicles during the communication process, the maximum distance $D_{\max}$ is given by

$$D_{\max} = \max\left(\frac{t^{\text{diff}}}{2}\left(\frac{g_{p,q}}{(T^j_{m+1} - T^j_m)} + \frac{g_{p,q}}{(T^k_n - T^k_{n-1})}\right), R_{\max}\right), \tag{10}$$

Algorithm 2: Relay Pair Matching

Input: Relay map $\mathbf{A}$
Output: Matching assignment $\mathbf{M}$

Function `Match(`i`)`**:**
  for *each* q_j^{in} *in* $\mathbf{Q}^{\mathbf{in}}$ *(in reverse order)* **do**
    if $e_{i,j} = 1$ *and* $\lambda_j = 0$ **then**
      $\lambda_j \leftarrow 1$;
      if q_j^{in} *not matched or* `Match(`$\mathbf{M}(j)$`)` **then**
        $\mathbf{M}(i) \leftarrow j$;
        return True;
      end
    end
  end
  return False;
initialization;
for *each* q_i^{out} *in* $\mathbf{Q}^{\mathbf{out}}$ **do**
  if q_i^{out} *not matched* **then**
    Set $\Lambda = \{\lambda_1, \cdots, \lambda_{|Q^{in}|}\}$ to zeros ;
    `Match(`i`)`
  end
end

where $g_{p,q}$ indicates the length of the road segment connecting node n_p and node n_q.

The second case is that the two vehicles drive in the same direction, as route i and route j in Fig. 2 show. Compared with the first case, the estimation of potential transmission amount is different, due to the difference in relative distance between vehicles. In this case, the meeting time of two vehicles is given by

$$t^{\text{same}} = \max\left(\min\left(T^i_{l+1}, T^j_{m+1}\right) - \max\left(T^i_l, T^j_m\right), 0\right). \tag{11}$$

Generally, the relative distance between two vehicles driving in the same direction is relatively stable during the communication. Therefore, the potential transmission amount can be calculated as follows,

$$\hat{E} = t^{\text{same}} W\log_2(1+\eta D^{-4}_{\text{init}}), \tag{12}$$

where the initial relative distance between two vehicles can be obtained by

$$D_{\text{init}} = \frac{\left|T^i_l - T^j_m\right|}{2}\left(\frac{g_{p,q}}{(T^i_{l+1} - T^i_l)} + \frac{g_{p,q}}{(T^j_{m+1} - T^j_m)}\right). \tag{13}$$

The relay process between two vehicles has greater success probability with the larger $\hat{E}$, which means that they should be given higher priority when making relay assignment. Considering the relay failure caused by a leaving vehicle being unmatched till the maximum waiting time, the leaving vehicle with longer waiting time should also be given higher relay priority. Thus, the relay weight ξ considering both the potential transmission amount and the waiting time is given by

$$\xi = \alpha\frac{w_i^{out}}{w_{\max}} + \beta\frac{\hat{E}}{B_i}, \tag{14}$$

where α and β denote the waiting time factor and the transmission amount factor, respectively. Vehicle q_i^{out} and vehicle

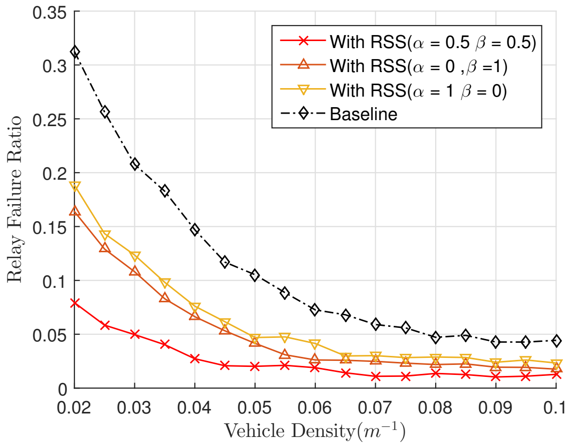

Fig. 4. The relay failure ratio and vehicle density with different relay selection strategy.

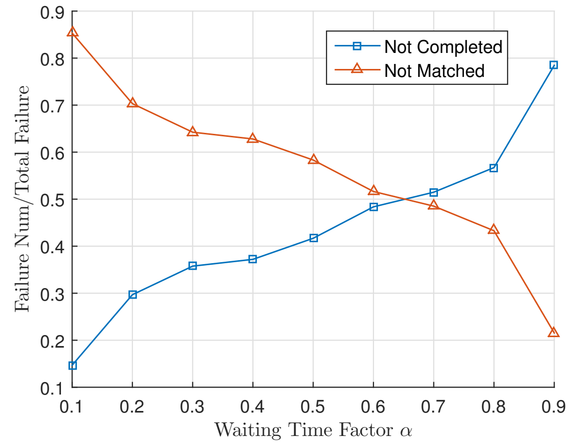

Fig. 5. The failure number percentage caused by different reasons with different waiting time factor.

q_j^{in} are connected with an edge $e_{i,j}$ if their relay weight meets the requirement, as Fig. 3 shows. Therefore, Q_{out} and Q_{in} are connected into a bipartite graph $\mathbf{A} = \left(\mathbf{Q^{out}}, \mathbf{Q^{in}}, \mathbf{E}\right)$, every edge of which connects a member in Q_{out} to one in Q_{in}.

B. Relay Pair Matching

Based on the relay map constructed by Algorithm 1, the optimization problem in (6) can be reformulated into the maximum matching problem as follows

$$\begin{aligned} \max_{\mathbf{M}} \quad & \sum_{i=1}^{\left|\mathbf{Q^{out}}\right|} e_{i,\mathbf{M}(i)} \\ \text{s.t.} \quad & \mathbf{M} \subset \mathbf{A}. \end{aligned} \tag{15}$$

The classic Hungarian algorithm for solving the maximum bigraph matching problem is employed with special adjustment for the relay selection problem. The detailed RPM algorithm using depth first search (DFS) is shown in Algorithm 2. In order to match the vehicles with longer waiting time first, the traversing order is specially condidered based on the characteristic of DFS. For q_i^{out} in $\mathbf{Q^{out}}$, the *forward* order is employed due to that the leaving vehicles being visited first have higher priority to be matched. The $1 \times \left|Q^{in}\right|$ vector Λ is employed to mark whether the incoming vehicles have been visited. In the recursion process of DFS, the matching results of the incoming vehicles that are visited later are returned first from *Match* function, which means that the incoming vehicles being visited later have higher priority to be matched. Therefore, in the *Match* function, q_j^{in} in $\mathbf{Q^{in}}$ is traversed in *reverse* order to match the incoming vehicles with longer waiting time. The matching assignment $\mathbf{M}$ selects incoming vehicles for the leaving vehicles to relay data, e.g., the matched pair $\mathbf{M}(i) = j + 1$ means that the i-th leaving vehicle relay its data to the j-th incoming vehicle, as shown in Fig. 3.

IV. Performance Evaluation

The efficiency of our proposed RSS are evaluated by simulations in terms of multiple performance metrics including relay failure ratio, retrieval throughput, and RSU consumption ratio, as defined in [16]. For the simulation setup, we select an area of size 680m $\times$ 350m from the map of Beijing using OpenStreetMap. Based on the realistic road network, the vehicular traffic is then generated by SUMO, an open-source simulator for road traffic simulation. In order to compare the proposed RSS with the baseline, the basic communication parameters used in our experiment are set to the same as those in [16].

As shown in Fig. 4, when the vehicle density increases, the relay failure ratio decreases since more vehicles provide more choices for data relay. The performance gain by the RSS is greater in the case of lower vehicle density, which indicates that the proposed RSS improves the robustness of the IVS system in low vehicle density scenarios. For all the three cases of waiting time factor and transmission amount factor settings, the relay failure ratio of the RSS is lower than that of the baseline. In particular, $\alpha = 0, \beta = 1$ and $\alpha = 1, \beta = 0$ represent the case of only considering transmission amount and the case of only considering waiting time, respectively. Compared with those extreme cases, the case of both considering transmission amount and waiting time, i.e., $\alpha = 0.5, \beta = 0.5$, has the lowest relay failure ratio, due to the factors leading to the relay failure are weighed overall. The effects of the waiting time factor and the transmission amount factor are also illustrated in Fig. 5. With the increasing of the waiting time factor, vehicles can be assigned with high priority to relay data even with low relay success probability, which reduces the occurrence of not attempting to transmit data even when deleted from the waiting queue. However, increasing the waiting time factor also brings more transmission failue due to the risky assignment. Therefore, as we can see in Fig. 5, with the increasing of the waiting time factor, the relay failure number caused by the leaving vehicles being unmatched decreases while the one caused by the incompleted transmission process increases.

The retrieval throughput gain is shown in Fig. 6. It can be observed that the IVS system has higher throughput compared with the in-vehicle storage disabled scheme due to the frequency spatial reuse by V2V links. Furthermore, with the RSS employed, the IVS system can achieve higher throughput by

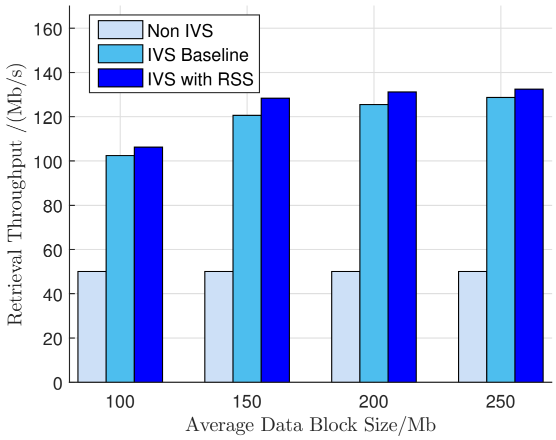

Fig. 6. The retrieval throughput and the average data block size with different data storage scheme.

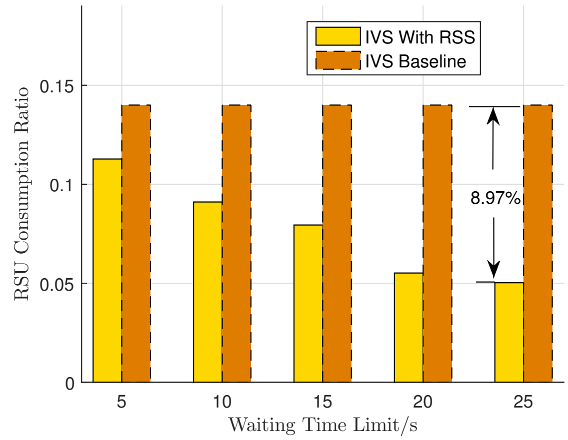

Fig. 7. The RSU consumption ratio and the maximum waiting time with different relay selection strategy.

reducing the data loss caused by relay failure. The impact of the waiting time limit is demonstrated in Fig. 7. With the increasing of the waiting time limit, the leaving vehicles can be assigned to more appropriate relay vehicles during a longer waiting time, which reduces the RSU consumption of resending the lost data. Therefore, the RSU consumption ratio can be reduced by increasing the waiting time limit, e.g., the RSU consumption ratio is reduced by 8.97% when the waiting time limit is set to 25 seconds.

V. Conclusion

In this paper, we proposed an RSS for the data relay stage of the in-vehicle storage system in vehicular networks. In order to cope with the relay failure, the proposed RSS first evaluated the relay success probability using the RMC algorithm, based on which, the relay selection problem was formulated into a maximum bigraph matching problem and solved by the proposed RPM algorithm. The performance of the RSS has been evaluated by a simulator consisting of realistic road network and vehicular traffic. The simulation results showed that the RSS performed better than the baseline in terms of relay failure ratio, retrieval throughput and RSU consumption ratio.

References

[1] M. C. *et al.*, "The Cityscapes Dataset for Semantic Urban Scene Understanding," in *Proc. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, Jun. 2016.

[2] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision Meets Robotics: The KITTI Dataset," *The International Journal of Robotics Research*, vol. 32, no. 11, pp. 1231–1237, Sep. 2013.

[3] T. Scharwächter, M. Enzweiler, U. Franke, and S. Roth, "Efficient Multi-Cue Scene Segmentation," in *Proc. German Conference on Pattern Recognition*, Saarbrücken, Germany, Sep. 2013.

[4] J. Jiao, "Machine Learning Assisted High-Definition Map Creation," in *Proc. 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC)*, Tokyo, Japan, Jul. 2018.

[5] J. Z. *et al.*, "Making Bertha Drive—An Autonomous Journey on a Historic Route," *IEEE Intelligent Transportation Systems Magazine*, vol. 6, no. 2, pp. 8–20, Summer 2014.

[6] Q. Li, L. Chen, M. Li, S. Shaw, and A. Nüchter, "A Sensor-Fusion Drivable-Region and Lane-Detection System for Autonomous Vehicle Navigation in Challenging Road Scenarios," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 2, pp. 540–555, Feb. 2014.

[7] N. Fairfield and C. Urmson, "Traffic Light Mapping and Detection," in *Proc. 2011 IEEE International Conference on Robotics and Automation*, Shanghai, China, May 2011.

[8] J. Pääkkönen, C. Hollanti, and O. Tirkkonen, "Device-to-Device Data Storage for Mobile Cellular Systems," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Atlanta, GA, USA, Dec. 2013.

[9] E. Altman, K. Avrachenkov, and J. Goseling, "Distributed Storage in the Plane," in *Proc. IFIP Networking Conference*, Trondheim, Norway, Jun. 2014.

[10] R. Ding, T. Wang, L. Song, Z. Han, and J. Wu, "Roadside-Unit Caching in Vehicular Ad Hoc Networks for Efficient Popular Content Delivery," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, New Orleans, LA, USA, Mar. 2015.

[11] L. Idir, S. Paris, and F. Naït-Abdesselam, "Optimal Caching of Encoded Data for Content Distribution in Vehicular Networks," in *Proc. IEEE International Conference on Communication Workshop (ICCW)*, London, UK, Jun. 2015.

[12] S. Zhang, N. Zhang, X. Fang, P. Yang, and X. S. Shen, "Cost-Effective Vehicular Network Planning with Cache-Enabled Green Roadside Units," in *Proc. IEEE International Conference on Communications (ICC)*, Paris, France, May 2017.

[13] N. Kumar and J.-H. Lee, "Peer-to-Peer Cooperative Caching for Data Dissemination in Urban Vehicular Communications," *IEEE Systems Journal*, vol. 8, no. 4, pp. 1136–1144, Dec. 2014.

[14] W. Zhao, Y. Qin, D. Gao, C. H. Foh, and H.-C. Chao, "An Efficient Cache Strategy in Information Centric Networking Vehicle-to-Vehicle Scenario," *IEEE Access*, vol. 5, pp. 12 657–12 667, 2017.

[15] B. Hu, L. Fang, X. Cheng, and L. Yang, "Vehicle-to-Vehicle Distributed Storage in Vehicular Networks," in *Proc. IEEE International Conference on Communications (ICC)*, Kansas City, MO, USA, May 2018.

[16] B. Hu, L. Fang, X. Cheng, and L. Yang, "Relay-and-Repair Based in-Vehicle Storage (R^2IVS) System in Vehicular Networks," in *Proc. 2018 IEEE Global Communications Conference (GLOBECOM)*, Abu Dhabi, United Arab Emirates, Dec. 2018.

[17] B. Hu, L. Fang, X. Cheng, and L. Yang, "In-Vehicle Caching (IV-Cache) Via Dynamic Distributed Storage Relay (D^2SR) in Vehicular Networks," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 1, pp. 843–855, Jan. 2019.

[18] J. Lemieux and Y. Ma, "Vehicle Speed Prediction Using Deep Learning," in *Proc. 2015 IEEE Vehicle Power and Propulsion Conference (VPPC)*, Montreal, QC, Canada, Oct. 2015.

[19] V. S. Tiwari, S. Chaturvedi, and A. Arya, "Route Prediction Using Trip Observations and Map Matching," in *Proc. 2013 3rd IEEE International Advance Computing Conference (IACC)*, Ghaziabad, India, Feb. 2013.

[20] G. Xue, Z. Li, H. Zhu, and Y. Liu, "Traffic-Known Urban Vehicular Route Prediction Based on Partial Mobility Patterns," in *Proc. 2009 15th International Conference on Parallel and Distributed Systems*, Shenzhen, China, Dec. 2009.