Master's Thesis
석사 학위논문

A Data Dissemination System towards Improving Channel Efficiency

Byungjin Ko (고 병 진 高 秉 鎭)

Department of Information and Communication Engineering

정보통신융합전공

DGIST

2016

Master's Thesis
석사 학위논문

A Data Dissemination System towards Improving Channel Efficiency

Byungjin Ko (고 병 진 高 秉 鎭)

Department of Information and Communication Engineering

정보통신융합전공

DGIST

2016

A Data Dissemination System towards Improving Channel Efficiency

Advisor: Professor Sang Hyuk Son

Advisor: Doctor Haengju Lee

Co-Advisor: Doctor Joonwoo Son

by

Byungjin Ko

Department of Information and Communication Engineering

DGIST

A thesis submitted to the faculty of DGIST in partial fulfillment of the requirements for the degree of Master of Science in the Department of Information and Communication Engineering. The study was conducted in accordance with Code of Research Ethics[1]

12 . 02 . 2015

Approved by

Professor (Advisor)	Sang Hyuk Son	(Signature)
Doctor (Advisor)	Haengju Lee	(Signature)
Doctor (Co-Advisor)	Joonwoo Son	(Signature)

[1] Declaration of Ethical Conduct in Research: I, as a graduate student of DGIST, hereby declare that I have not committed any acts that may damage the credibility of my research. These include, but are not limited to: falsification, thesis written by someone else, distortion of research findings or plagiarism. I affirm that my thesis contains honest conclusions based on my own careful research under the guidance of my thesis advisor.

A Data Dissemination System towards Improving Channel Efficiency

Byungjin Ko

Accepted in partial fulfillment of the requirements for the degree of Master of Science.

12 . 02 . 2015

Head of Committee ______________(인)
Prof. Sang Hyuk Son

Committee Member Haengju Lee ______________(인)
Dr. Haengju Lee

Committee Member ______________(인)
Dr. Joonwoo Son

MS/IC
201422001

고 병 진. Byungjin Ko. A Data Dissemination System towards Improving Channel Efficiency. Department of Information and Communication Engineering. 2016. 43p. Advisors Prof. Sang Hyuk Son, Advisors Dr. Haengju Lee, Co-Advisor Dr. Joonwoo Son.

ABSTRACT

Recent advancement of wireless communication technology in vehicular ad-hoc networks (VANETs) is expected to make various driving-related applications, which are envisioned by people, come true in near future. Data dissemination plays an important role to enable the emerging applications. Nevertheless, due to the limited bandwidth resource and the dynamic traffic feature, it is challenging to provide efficient data services in VANETs. In this work, we present a data dissemination system based on a hybrid of infrastructure-to-vehicle (I2V) and vehicle-to-vehicle (V2V) communications and a cooperation among multiple roadside units (RSUs) on a two-way roadway. The primary objective is to best explore the channel efficiency for both I2V and V2V communications and offload workloads of RSUs so that the system performance on data service can be maximized. To this end, we propose and integrate three approaches to enhance channel utilization and offload the workloads of RSUs. Specifically, in RSU's coverage, data items are disseminated via the hybrid of I2V and V2V communications. For the vehicles travelling out of the RSU's coverage, certain vehicles will be designated by RSUs to provide data services via V2V communication, which are called server-vehicles. Finally, RSUs located at different spots are designed to cooperate with each other for I2V data dissemination and a server-vehicle designation. An extensive performance evaluation demonstrates that the algorithm which combines the above three approaches can achieve higher service ratio, lower service delay, and higher fairness compared with the most competitive solution in the literature.

Keywords: data dissemination, RSU cooperation, scheduling, hybrid of I2V and V2V communications

Contents

List of tables

List of figures

I. Introduction

Efficient data dissemination is one of the most critical issues in vehicular ad-hoc networks (VANETs) to enable a variety of emerging intelligent transportation systems, such as the applications to enhance vehicle safety, transportation efficiency, and passenger comfortableness, etc. [1]. According to [1] and [2], it is expected to provide the services such as the parking lot booking, restaurant information (e.g., distance, food price, and menu), road sign notifications, map download, etc. in VANETs. Dedicated short range communication (DSRC) is one of the most promising wireless communication technologies, which is dedicated to enable both infrastructure-to-vehicle (I2V) and vehicle-to-vehicle (V2V) communications in VANETs. In general, DSRC is a suite of protocols including IEEE 802.11p, IEEE 1609.1/.2/.3/.4 [3], and SAE J2735 message set dictionary [4]. In order to build VANETs through DSRC, two kinds of devices are utilized: a roadside unit (RSU) and an on-board unit (OBU). The RSU is a fixed infrastructure installed along the road to provide data services. The OBU is mounted on a vehicle to enable V2V and I2V communications.

According to [5], the world vehicle population is estimated to reach 1.34 billion by 2016. Given the trend of the increasing number of vehicles, when the DSRC devices come into wide use, there will be tremendous demands on data access in VANETs to enable different applications, resulting in excessive channel access and increased workloads at RSUs. The performance of data services will be degraded dramatically due to the limited bandwidth resource when the system workload is getting heavier. Therefore, it is imperative to improve the channel efficiency on data services in VANETs.

Scheduling data dissemination in VANETs has been extensively studied in previous works. Relevant studies can be generally divided into single RSU based and multiple RSUs based data services. The single RSU based scheduling schemes only contemplate the data services for the vehicles within the RSU's coverage [6-8]. On the multiple RSUs based data service, the RSUs can cooperate to complete data services and further enhance the overall system performance. In such a scenario, each RSU may serve a part of the request [9-11]. Nevertheless, none of the previous work has considered the cooperation among RSUs by exploiting both the I2V and V2V communications in the coverage of each RSU, as well as V2V communications outside any RSU's coverage. In this work, we present a multiple RSUs based data dissemination system which considers providing the delay tolerant services by coordinating the hybrid of I2V and V2V communications in each RSU's coverage and designating proper sender vehicles for V2V data sharing out of RSUs' coverage via the cooperation of RSUs. The primary objectives are to enhance overall data service performance and offload workloads at RSUs.

Cooperative I2V and V2V communications within RSUs' coverage is a promising mechanism to improve the channel efficiency and balance the workloads. The service via pure I2V communication will not only increase the workloads at the RSUs, but also decrease the system scalability due to the limited service range of the RSU and the speed of vehicle. To deal with the observed problems, we consider designating the proper vehicles to assist data services so that they are able to serve the requests with their cached data items via V2V communication. However, it is challenging to design an efficient and cooperative data service mechanism via I2V and V2V communications. First, adjacent vehicles that can communicate with each other may suffer from interference when they try to exploit the same channel simultaneously to transmit or receive data items [12, 13]. Second, with single-radio OBUs, vehicles can only switch to either I2V or V2V mode at a time [14]. Third, current half duplex DSRC devices cannot transmit and receive data items simultaneously.

In addition to exploiting cooperative I2V and V2V communications within the RSU's coverage, it is expected to further enhance the channel efficiency and reduce the workloads by serving requests out of the RSU's coverage via V2V communication. In this work, we consider designating proper vehicles as 'server-vehicles', so that they can broadcast their cached data items to their neighbors via V2V communication. The server-vehicles can potentially serve those vehicles which are driving in the opposite direction when they are meeting somewhere along the two-way road. There are challenges to select the server-vehicles. First, RSUs need to select proper vehicles to improve data service performance by considering the cached contents in vehicles. Second, there is a trade-off between the number of server-vehicles and the V2V communication interference. That is, although more server-vehicles may give a better chance to serve more requests, excessive designating sender vehicles would dramatically increase the signal interference at the receiver, causing unsuccessful services. Third, there is another trade-off at the RSU between directly serving vehicles in its own coverage and indirectly serving other vehicles by designating a server vehicle. In other words, in order to let the server-vehicle retrieve and cache the data items required by vehicles from other RSUs, this RSU has to allocate certain bandwidth (or time slots) to serve the server-vehicle, and thus it may scarify the performance of serving the vehicles in its own coverage if the schedule is not well designed.

With the above analysis, we focus on the data dissemination on a two-way roadway where multiple RSUs cooperate in providing data services. This work is dedicated to investigating how to efficiently offload the workloads among multiple RSUs, designate server-vehicles, and schedule data dissemination via the hybrid of I2V and V2V communications, so that the system performance can be maximized in terms of improving the service ratio, reducing the service delay, and guaranteeing the fairness of data service.

The main contributions of this work are outlined as follows. First, we present a multi-RSU cooperated data service system where both the I2V and V2V communications inside RSUs' coverage and the V2V communication outside RSUs' coverage are exploited to enhance the data service performance. Second, we propose a solution which combines the following three approaches: the hybrid of I2V and V2V communications in RSU's coverage, the cooperation among RSUs, and the server-vehicle designation for data services out of RSUs' coverage. Third, we build the simulation model and give a comprehensive performance evaluation to demonstrate the superiority of the proposed solution.

The rest of this paper is organized as follows. Section II represents the related works. Section III illustrates the system model. A scheduling algorithm is proposed in section IV. In section V, we build the simulation model and give a performance evaluation. Finally, we conclude the thesis and discuss future work in section VI.

II. Related Work

Researches relative to the data dissemination in VANETs have been extensively studied in many scenarios, which are in terms of using pure I2V communication [8, 15], utilizing I2V and V2V communications [6, 16, 17], and co-working among neighboring RSUs [11, 18]. First of all, with regard to the pure I2V communication, K. Liu et al. present the study on real-time data services via I2V communication by considering both the time constraint of data dissemination and the freshness of data items [8]. In [15], maximum freedom last scheduling algorithm for downlinks from RSU to OBU is proposed; the service priority is determined by remaining dwell time of service channel, remaining transmission time, queueing delay, and maximum tolerable delay.

With respect to using I2V and V2V communications, K. Liu et al. present the data dissemination system using cooperative I2V and V2V communications, towards scalable, fair, and robust data service [6]. It proposes an on-line scheduling algorithm which is transforming the data dissemination problem to the maximum weighted independent set (MWIS) problem. J. Wang et al. present the cluster-based data sharing model, which is that a vehicle in a cluster transmits a data item to a vehicle in the opposite cluster via V2V communication [16]. Different time slots are assigned to the adjacent clusters to alleviate interference. RSU coordinates the length of the cluster based on the signal-to-interference-and-noise ratio (SINR) in V2V communications. (i.e., RSU divides the vehicles into different clusters.) Q. Wang et al. present the network-coding based joint V2I downlink and V2V communication system, which is that a timely information is relayed via V2V communication from a RSU to a vehicle which is out of RSU's coverage [17]. To this end, they derive the maximum throughput of the V2I downlink system considering additive white

Gaussian noise (AWGN) channels and Rayleigh fading channels with Doppler effects and discuss the maximum achievable amount of information which can be relayed from RSU to the vehicle.

In respect of the cooperation among multiple RSUs, in [18], transferring the delay tolerant request, which is not served via I2V communication, to the next RSU where the vehicle is heading is proposed in order to balance workloads among RSUs. The transferred requests can be served when the vehicle reaches the next RSU's coverage. Y. Gui and E. Chan propose a motion prediction based scheduling scheme, which enables cooperative work among a set of RSUs for scheduling multi-item requests by transferring the data requests to the neighbor RSU [11]. The transferred requests can be served at the next reaching RSU's coverage.

The problem investigated in our work is distinguished from the above studies in terms of the following aspects. First, we consider the maximal data service, which is delay tolerant, at a time via the hybrid of I2V and V2V communications. Second, RSUs cooperate with each other by transferring the unserved requests to designate server-vehicles which serve them out of RSU's coverage via V2V communications.

III. Data Dissemination System

The data dissemination system in a multi-RSU cooperation environment is represented in Figure 1. RSUs are installed along the road and interconnected via a wired backbone network so that they can share information and cooperatively provide data services [19]. OBUs are mounted on the vehicles to enable I2V and V2V communications. The dotted circle means RSU's coverage. The shadowed block represents the data item which has been retrieved and cached. The non-shadowed block indicates the outstanding request. The dashed bidirectional arrow between two vehicles represents that they can communicate with each other through V2V communication. The lined arrow represents data transmission via I2V or V2V communication according to a specific algorithm. Note that vehicles submit their requests only when they are in RSUs' coverage and each request corresponds to one data item.

Five DSRC channels are reserved in the system, including one control channel (CCH) and four service channels (SCH). The CCH is used to transmit basic safety messages (BSMs), probe vehicle data (PVD) messages, and WAVE service advertisement (WSA) messages in order to

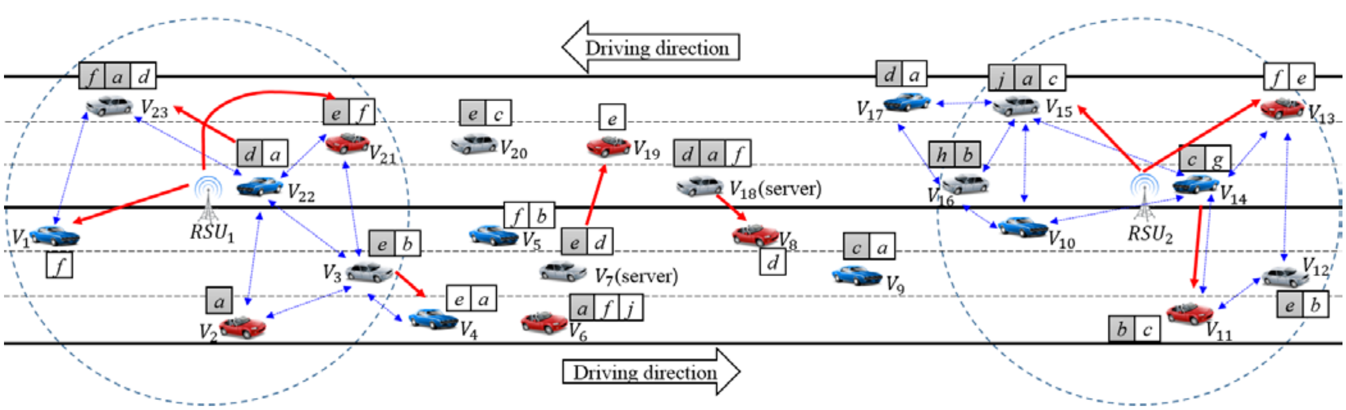


Figure 1. Data Dissemination System in a Multi-RSU Cooperation Environment

transmit or receive the information related to the auxiliary information for scheduling or the scheduling result which indicates the information of the sender, the receiver, and the data item to share. Two SCH are respectively used for I2V and V2V communications in the RSU's coverage. The other two SCH are respectively used depending on the server-vehicle's driving direction out of the RSU's coverage in order to provide data service for client-vehicles, which are not designated as a server-vehicle.

In the system, the data service is provided through two kinds of data dissemination schemes depending on the different service areas. One is the data dissemination via cooperative I2V and V2V communications within RSUs' coverage. The other is data dissemination through pure V2V communication out of RSUs' coverage. Detailed operations of the system are presented as follows.

A) Hybrid of I2V and V2V communications in RSU's coverage

Within RSUs' coverage, vehicles and RSUs periodically comply with the four phases, which is described in Figure 2, for the scheduling data dissemination.

- In the phase (1), every vehicle broadcasts the heartbeat message (BSM) to its neighboring vehicles through CCH. As a result, every vehicle can maintain a list of its neighboring vehicles.
- In the phase (2), every vehicle transmits a PVD message, which is defined in [4], including the list of the cached data items, the requiring data items, and information of neighboring vehicles to the RSU. The message is transmitted through CCH via V2I communication.
- In the phase (3), every RSU schedules data dissemination based on the received PVD messages from the vehicles. After that, the RSU broadcasts the WSA message to every

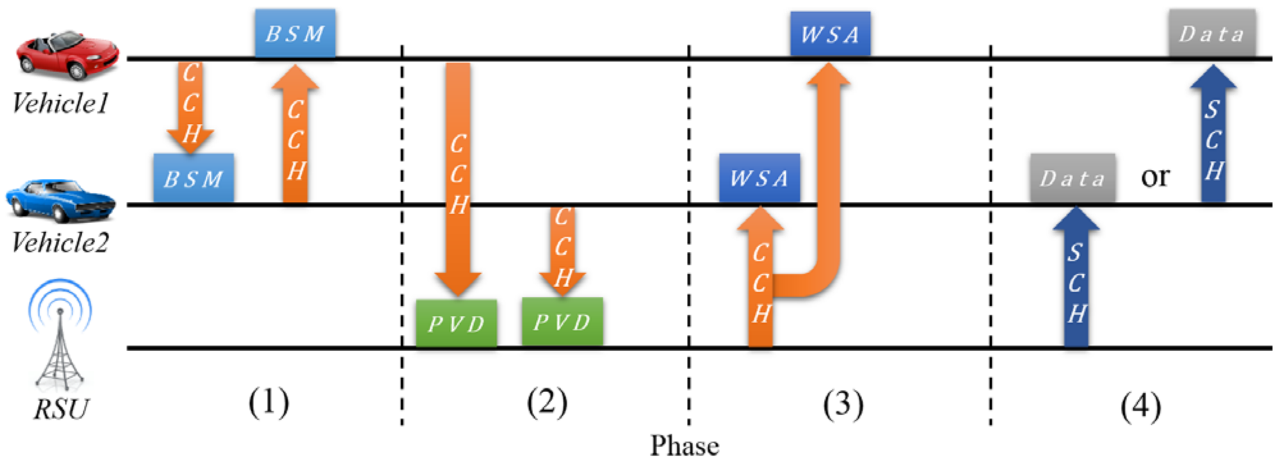


Figure 2. Phase for Scheduling Data Dissemination within RSU's Coverage

vehicle in its coverage through CCH. The message contains the determination including the sender vehicle IDs and the corresponding data IDs to be shared, as well as the receiver vehicle IDs and their communication modes, either V2V or I2V, to be set in the following phase.

- In the phase (4), according to the scheduling result, data items are simultaneously disseminated via the cooperative I2V and V2V communications, where V2V communications are constrained to one-hop. In addition to serving vehicles in its own coverage, every RSU monitors unserved requests submitted by the vehicles which are leaving the RSU's coverage and informs other RSUs where the vehicles are heading, so that other RSUs may prepare in advance for serving these outstanding requests.

The additional information such as the lists of the cached data items, the requiring data items, and the neighboring vehicles is not described in a PVD or a WSA message format. The size of messages depends on the number of vehicles within V2V or I2V communication coverage in our system. We verify that the available fields of the messages are enough to include the additional information and describe the possible number of vehicle IDs, which can be included in the PVD message, and the availability of using the WSA message in APPENDIX A.

As shown in Figure 2, RSUs schedule data dissemination through the above phases. In RSU_1's coverage, data items are transmitted through cooperative I2V and V2V communications. In addition, the unserved requests, which are submitted by the vehicles leaving RSU_1's coverage, are transferred to RSU_2 by RSU_1. On the other hand, in the current time slot, data items are mainly disseminated via V2V communication in RSU_2's coverage, because RSU_2 appoints V_{15} as a server-vehicle and broadcasts the data item f to V_{15}, anticipating that V_{15} can serve the transferred requests from other RSU.

B) Pure V2V communication out of RSU's coverage

In order to provide the data service out of RSUs' coverage with minimum transmission collisions at client-vehicles, RSUs are expected to designate proper server-vehicles at the pertinent timing. The server-vehicles are selected based on the information of transferred requests from other RSUs and vehicles' cache contents when the previous server-vehicle is expected to be moved more than V2V transmission range after the server-vehicle designation. For the data service out of RSU's coverage, all server-vehicles and client-vehicles follow the three phases in Figure 3. Note that, a server-vehicle provides the data service only for a client-vehicle driving in the opposite direction.

- In phase (1), a server-vehicle broadcasts a WSA message, which is including an identifier of a server-vehicle, the list of cache contents, and SCH information for the service, to client-vehicles via CCH when the server-vehicle does not provide data service.

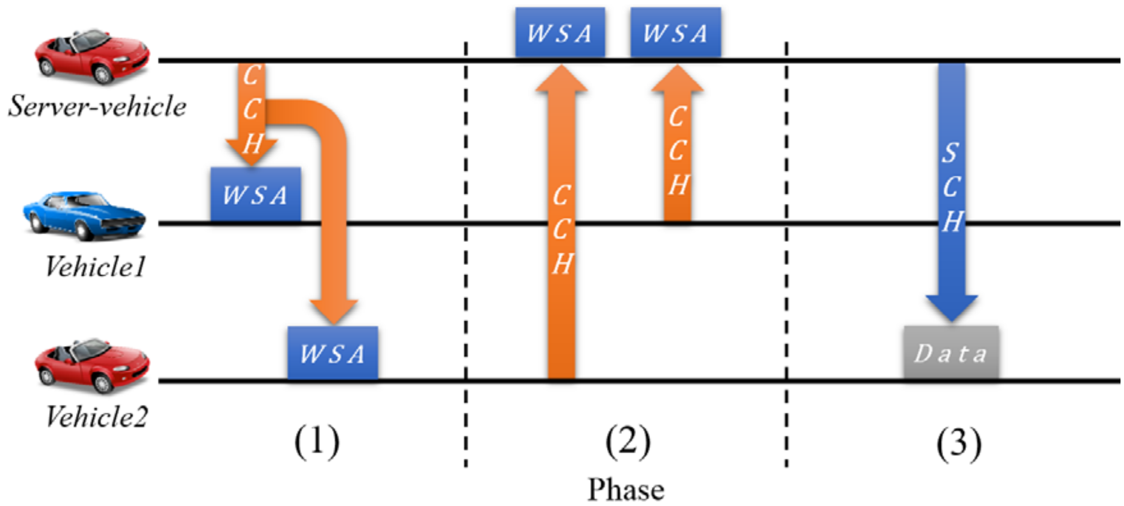


Figure 3. Phases for Data Service out of RSU's Coverage

- In phase (2), a client-vehicle transmits a WSA including an identifier of request to the server-vehicle as a response when the client-vehicle is aware that the server-vehicle can serve one of the outstanding requests. After that, the client-vehicle tunes to the SCH that the server-vehicle uses to provide the data service. If a server-vehicle detects the other server-vehicle which is driving in the opposite direction and caching the data items matching up to the outstanding requests, it follows the same steps. Note that, the server-vehicle expecting to receive the data item will retune to the other SCH to provide data service for a client-vehicle if the server-vehicle notices that the other server-vehicle is providing the data service for other client-vehicle.
- In phase (3), the server-vehicle broadcasts the data item, which is corresponding to the first received WSA from the client-vehicle, through the SCH.

Note that a server-vehicle cannot transmit and receive data items at the same time. For instance, in this time slot, V_7 is transmitting *e*, so it cannot receive *d* from V_{18} even though *d* is required by V_7.

IV. Proposed Algorithm

In order to provide the data service towards improving channel efficiency and implementing workload offload, we propose a scheduling algorithm called *Less Conflict Selection and Server-vehicle Designation (LCSD)*, which coordinates I2V and V2V communications in the RSU's coverage by transforming the scheduling problem to the maximum independent set (MIS) problem and designates a server-vehicle through the RSUs' cooperation. As a result of *LCSD*, a service list is generated, which consists of MIS and an identifier of data item to transmit to a designated server-vehicle. In order to briefly represent the data sharing between vehicles and between a RSU and a vehicle, the RSU and the vehicle are respectively abbreviated to *R* and *V*. Moreover, transmitting a data item to a receiver is indicated with "sender", "data item", and "receiver" in the order named. To name but a few, R_1fV_{21} means R_1 transmits *f* to V_{21} and V_3eV_4 signifies V_3 sends *e* to V_4.

A) Problem transformation

In order to solve MIS problem, we adopt a greedy method. The detailed rationality of graph constructing can be referred from [6]. Note that the major difference of this work in problem transformation is that we consider different issues when constructing the weighted graph.

Services of data dissemination using I2V and V2V communications can be classified as guaranteed service, unfeasible service, and potential service. Firstly, the guaranteed service is that the service should be surely provided for the vehicle by RSU in any given situation. In other words, the data service should be guaranteed, which is the data transmission from RSU to the designated server-vehicle, and surely included in the service list. For example, I2V communication to transmit the data item, corresponding to the transferred request, from RSU to the designated server-vehicle ($R_2 fV_{15}$ in Fig. 1) is included in the service list.

Secondly, the unfeasible service includes two cases: the first one is I2V communication, which disturbs the transmission of data item to serve the transferred request from RSU to the designated server-vehicle (e.g., R_2cV_{11}, R_2bV_{12}, R_2eV_{13}, R_2gV_{14}, *and* R_2cV_{15} when designated server-vehicle is V_{15} and the data item to be transmitted from RSU to the designated one is *f* in Figure 1), and the second one is V2V communication between the designated server-vehicle and its neighbor. (e.g., $V_{15}aV_{17}$, $V_{14}cV_{15}$ in Figure 1) They are excluded in the service list because they interrupt the data transmission from the RSU to the designated server-vehicle.

Finally, the potential service is related to a request which may be served either from the RSU via I2V communication or from the neighboring vehicle via V2V communication without interfering the data transmission from RSU to the designated server-vehicle. For example, if receiving the same data item via I2V communication does not disturb the data transmission from RSU to the designated server-vehicle, it is a potential service. (e.g., R_2fV_{13} and every feasible I2V and V2V communications in RSU_1's coverage in Figure 1)

Note that *LCSD* only considers potential services and guaranteed services.

1) Graph construction

The RSU coordinates I2V and V2V communications by transforming the problem of the scheduling data dissemination to a graph model. The graph consists of a set of vertices and edges, which is a non-directional graph.

In a graph, only the potential service is expressed as a vertex. It is represented by a circle with a label which is associated with the sender, the data item to be disseminated, and the receiver information. The sender or the receiver can be either a RSU or a vehicle. Therefore the vertices can represent both I2V and V2V services in the transmission range of a RSU.

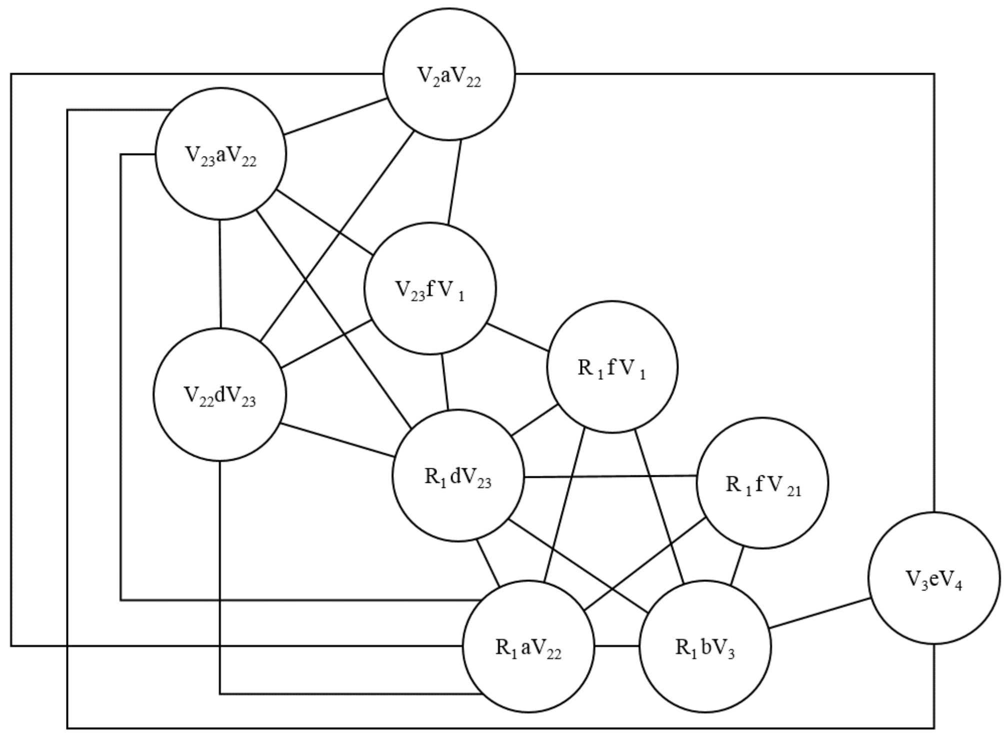
Each edge in our graph represents a conflicting relationship between two vertices. If there are two of potential services not to be operated at the same time due to certain constraints, these services are conflicted. Two vertices related to these conflicted services are connected via an edge. The number of edges connected to a vertex means the conflict degree. Detailed constraints and conflicting relationships are defined as follows along with examples shown in Figure 1.

- If two vertices both represent I2V communications and they represent that the RSUs broadcast different data items, they are in conflict with each other, because RSU cannot broadcast more than one data item at a time. For instance, R_1fV_1 and R_1bV_3 have the conflicting relationship with each other.
- If one vertex represents I2V communication and the other one represents V2V communication and the receiver of the former vertex is same with the receiver or transmitter of the later vertex, then they are in conflict with each other. That is because OBUs cannot receive and transmit the data item at the same time. In addition, the OBUs can only in I2V or V2V mode at a time. For instance, R_1dV_{23} has conflicting relationships with $V_{22}dV_{23}$ and $V_{23}fV_1$ respectively.
- If two vertices both represent transmitting different data items via V2V communications and the senders of both vertices are the same vehicle, then they are in conflict with each other, because OBUs can transmit only one data item at a time. For instance, $V_{23}aV_{22}$ and

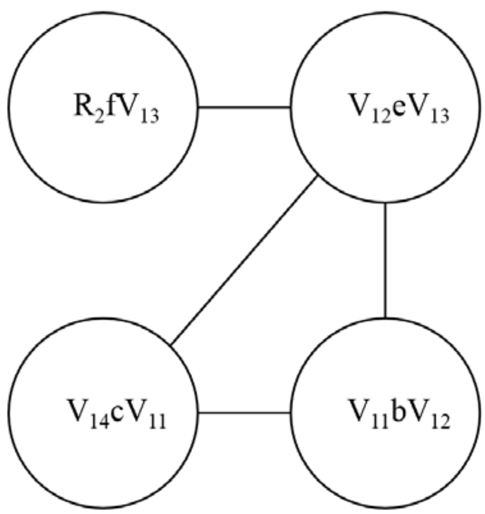
$V_{23}fV_1$ have the conflicting relationship with each other.

- If two vertices both represent V2V communications and a sender of one vertex and a receiver of the other are the same vehicle, then they are in conflict with each other, because OBUs cannot be a sender and a receiver at the same time. For instance, $V_{22}dV_{23}$ and $V_{23}aV_{22}$ have the conflicting relationship each other.
- If two vertices both represent V2V communications and a receiver is a neighboring vehicle of both the senders, then they are in conflict with each other, because data collision occurs at the receiver. For instance, V_2aV_{22} has the conflicting relationship with $V_{23}fV_1$.

Figure 4 shows how to transform potential operations to vertices. (a) and (b) are constructed based on the received WSA message from the vehicles in RSU_1's and RSU_2's coverage respectively.



(a) The graph constructed based on the potential service in *RSU_1*'s coverage



(b) The graph constructed based on the potential service in *RSU_2*'s coverage

Figure 4. An example of problem transformation

2) Weight assignment and algorithm implementation

To serve as many requests as possible, it is expected to find the MIS of the graph. A greedy method is adopted to approximately solve MIS problem. The general idea of such a greedy method is recapitulated as follows. First, it selects a vertex with the maximum weight, and then it removes this vertex and its neighbors. It repeats this process on the remaining graph until no vertex remains. The selected vertices are included in the service list. *LCSD* intends to select a vertex which has a less number of neighbors because the less eliminated number of the vertex i's neighbors is, the more the potential services are remained. Furthermore, to that end, it is considered to eliminate a lot of the neighbors' edges for reducing service conflicts. Based on the above mentioned concepts, we assign the weight to each vertex by considering three factors. One is the number of vertex's edges, which is denoted by $|E_i|$, where $E_i = \{e_1, e_2, \ldots, e_{|NbrV_i|}\}$ is set of edges of the vertex i. $NbrV_i$ is the set of vertex i's neighbors, and $|NbrV_i|$ is the number of neighbors connecting to the vertex. This factor represents the number of vertices to be removed in selecting the vertex. The second factor is the average number of neighbors' edges $|E|$, i.e., AN_i, which is defined as

$$AN_i = \{nv|(\sum_{nv \in NbrV_i} E_{nv})/|E_i|\}. \quad (1)$$

AN_i implies the conflict degree of neighboring vertices. Since the selected vertex i and its neighbors are eliminated by greedy method, the more the value of AN_i is, the more each neighbor's edges is eliminated. Therefore, the conflict degree of the remained vertices in the graph will be generally decreased. The third factor is the sum of neighbors' AN, i.e., SNA_i, which is given by

$$SNA_i = \{nv| \sum_{nv \in NV_i} AN_{nv}\} . \quad (2)$$

This factor complements the incompleteness of greedy method in terms of not guaranteeing optimal solution. Therefore, *LCSD* considers the vertices which can be affected by selection of vertex *i* and may be chosen as independent set later. SNA_i means the conflict degree of each vertex which is connected to all neighbors of vertex *i*. The less value of this factor helps to select a vertex *i* towards less conflicts on the vertices connected to vertex *i*'s neighbors so that their neighbors can be less removed when they are selected as the set later.

Given the above three factors, the weight of vertex i, W_i is represented by

$$W_i = -E_i + AN_i - SNA_i \qquad (3)$$

The higher value of W_i represents the less amount of vertices to be eliminated by the greedy method. (i.e., $e_{R_1fV_1}$= 4, $AN_{R_1fV_1}$= 6, $SNA_{R_1fV_1}$= 20.2, $W_{R_1fV_1}$= - 18.2)

After the greedy method, the unserved requests transmitted by the vehicles which are leaving the RSU's coverage are transferred to the next RSU, as the cooperation among multiple RSUs. After that, the RSU designates a proper server-vehicle based on the transferred request.

B) Server-vehicle designation

After receiving the transferred requests from the neighboring RSUs at every scheduling period, RSU accumulates the requests in the service queue for a server-vehicle designation. RSUs are supposed to designate server-vehicles whenever the latest server-vehicle is assumed to be passed the distance of V2V range based on the average speed of vehicles within their coverage. To designate a server-vehicle, the RSU nominates a set of candidate-server-vehicles, which are leaving the RSU's coverage, and assigns a weight to each candidate. After that, the RSU designates a server-vehicle, which has the highest weight among them.

Two factors are considered to assign the weight of server-vehicles. One is the number of cached data items which can serve the transferred requests. The other is the total number of cached data items. The weights of candidate-server-vehicles are denoted by $CW(t)=\{cw_1, cw_2,\ldots, cw_{|cv(t)|}\}$, where $|cv(t)|$ is the total number of candidate-server-vehicles at time t. Each weight is represented by

$$cw_i = \alpha|CCT_i(t)| + \beta|CD_i(t)|, \quad (4)$$

where α and β are tuning parameters to differently weight between the available service of transferred requests and the service capacity out of RSU's coverage. To obtain cw_i, first, the RSU measures the service capacity of candidate-server-vehicle by checking the set of cached data items of each candidate-server-vehicle, which is denoted by $CD_i(t) = \{cd_1, cd_2,\ldots,cd_{|CD_i(t)|}\}$, where $|CD_i(t)|$ is the total number of cached data items of the vehicle i at time t. Second, the RSU inspects the available services, the cached data items corresponding to the transferred requests, which is denoted by

$$CCT_i(t) = TR(t) \cap CD_i(t) \quad (5)$$

where the RSU can obtain the set of transferred requests from neighboring RSUs, which is denoted by $TR = \{tr_1, tr_2,\ldots, tr_{|TR|}\}$, where $|TR|$ is the total number of transferred requests. The number of $CCT_i(t)$'s elements is denoted by $|CCT_i(t)|$. (i.e., $TR = \{j, f, f, b, a\}$, $CCT_{V_{15}}(t) = \{a, j\}$, $CCT_{V_{16}}(t) = \{h, d\}$, $cw_{V_{15}} = 2\alpha+2\beta$, $cw_{V_{16}} = \alpha+2\beta$)

After designating the server-vehicle, RSU selects the hottest requested data item among the transferred requests and transmits it to the selected vehicle if it has not yet cached this data item. If the vehicle has already cached the data item, then the RSU transmits the second hottest data item. After designating the server-vehicle, the RSU resets the queue and then it accumulates the transferred requests until appointing the next server-vehicle.

V. Performance Evaluation

The simulation model is built based on the system architecture presented in Section III, and it is implemented by CSIM 19 [20]. We evaluate the performance with two types of traffic model, the real traffic model and the Greenshield's model, in order to show that *LCSD* is scalable in any traffic condition, either light or heavy traffic. The road environment in simulation is a highway, but it can be extended to a city street where the straight is long.

As the real traffic information, we utilize the next generation simulation (NGSIM) data sets, which capture vehicle trajectories in every 0.1 second. One is U.S. Highway 101 (US-101) in Los Angeles, California collected between 7:50 a.m. and 8:05 a.m. on June 15, 2005 [21]. The other is Interstate 80 (I-80) in Emeryville (San Francisco), California collected between 4:00 p.m. and 4:15 p.m. on April 13, 2005 [22]. In each direction, there are 6 lanes on the road, except an on-ramp. Because each data set contains the vehicle trajectory information regarding one way direction, we

TABLE 1. Simulation statics in scenario US-101

Time Period	Flow (vehicle / hour)	Mean Speed (m / hour)
7:50 am – 7:55 am	9156	31.60
7:55am – 8:00 am	8820	26.60
8:00 am – 8:05 am	7560	25.59
Average	8612	28.06

use the traffic mirroring which uses the same information to the opposite direction. The detailed statistics are summarized in TABLE 1 and 2.

Greenshield's model is one of the traffic models, which assumes a linear relationship between the vehicle speed and the traffic density [23]. The vehicle speed (v) is represented by $v = V^f - \frac{V^f}{K^j} \cdot k$, where V^f is the free flow speed (i.e. the maximum speed of vehicles) and K^j is the jam density (i.e. the density leading to zero velocity of vehicles). k is the current density of vehicles on each lane. In each direction, there are 3 lanes on the road, except an on-ramp and an auxiliary lane. The default parameters are set as follows. The free flow speeds of the first, second and third lanes are respectively set to v_1^f=120 $^{km}/_{h}$, v_2^f=110 $^{km}/_{h}$, and v_3^f=100 $^{km}/_{h}$ and K^j is set to 100 $^{vehicles}/_{km}$. We evaluate the performance in a wide range of traffic workloads, and the detailed statistics are summarized in TABLE 3. The larger traffic volume level corresponds to the heavier traffic workload.

TABLE 2. Simulation statics in scenario I-80

Time Period	Flow (vehicle / hour)	Mean Speed (m / hour)
4:00 pm – 4:05 pm	6612	23.44
4:05 pm – 4:10 pm	8364	22.13
4:10 pm – 4:15 pm	7800	20.95
Average	7592	22.11

TABLE 3. Simulation statics under different traffic scenarios (both directions)

Traffic Volume Level	Mean Arrival Rate (vehicle / hour)			Mean Speed (km / hour)		
	First Lanes	Second Lanes	Third Lanes	First Lanes	Second Lanes	Third Lanes
1	1300	1200	1100	105.30	96.48	86.96
2	1600	1500	1400	99.20	91.99	84.02
3	1900	1800	1700	96.45	87.07	80.15
4	2200	2100	2000	92.06	78.91	73.11
5	2500	2400	2300	88.14	75.58	64.29

The distance between the two RSUs is set to 4 km. The RSU’s communication radius is set to 300 m and the OBU’s communication range is set to 150 m. The OBU can transmit PVD message to RSU through CCH in RSU’s coverage. The scheduling period is 1sec, which is reasonable because the maximum transmit time for a full-size MAC service data unit (2312 octets) is approximately 6.5ms [24] and 1 sec is enough for both computation and transmission time. The specific description of the computation time of algorithm can be referred to APPENDIX B. The database consists of 200 data items: map data of different spots, road condition, location-based advertising, etc.. It is a reasonable size because some data items are dependent on locations and different each other although they are used for the same content (e.g., map data item A, B, and C, which represent the dissimilar areas, are categorized as the same content but they are different.). Each vehicle generates 1 to 10 requests within every RSU’s coverage, and the data access pattern follows the Zipf distribution [25] with the parameter θ=0.4. The threshold to designate candidate-server-vehicles is set to 255 meters away from the RSU. The tuning parameters α and β are set to 0.4 and 0.6 respectively.

We compare the proposed algorithm with MRF (Most Requested First) [26] and SFR (Scalability, Fairness, and Robustness) [6]. MRF broadcasts the data items to serve the hottest data item via I2V communication. SFR schedules data dissemination in a single RSU environment via cooperative I2V and V2V communications. We evaluate the algorithm performance with the following three metrics.

- Total service ratio: This metric is to evaluate the channel efficiency and the workload offload. The total data service ratio is determined by two parameters: the number of total submitted requests (n_t) and the number of serviced requests (n_s). It is computed by this equation: $\frac{n_s}{n_t}$. The higher service ratio implies the more efficient channel utilization and the better performance on workload offload.
- Average service time: This metric is to evaluate the service quality in terms of service waiting time. For a request, its service time is from the instant when the request is submitted to the time when the corresponding data item is retrieved. The average service time of all the successfully served requests measures the responsiveness of the system. The shorter average service time means the better service quality regarding the response time.
- Cold data service ratio: This metric is to evaluate the fairness of service, in terms of that the pattern of data access follows the Zipf distribution. The cold data is a data item which is infrequently accessed in a database, which is defined as having a data access probability

less than $\frac{1}{|D|}$, where $|D|$ is the number of data items in the data base. The cold data service ratio is determined by two parameters: the total number of generated requests for cold data items (n_c) and the number of served requests among them (n_{sc}). To be specific, the cold data service ratio is calculated with this equation: $\frac{n_{sc}}{n_c}$. The higher ratio implies the fairer services are provided for the vehicles.

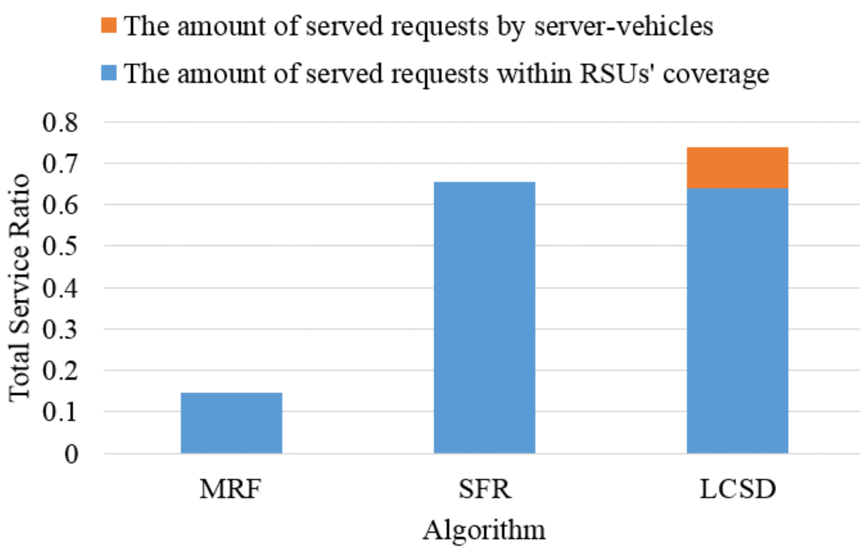


Figure 5. Total Service Ratio in Scenario US-101

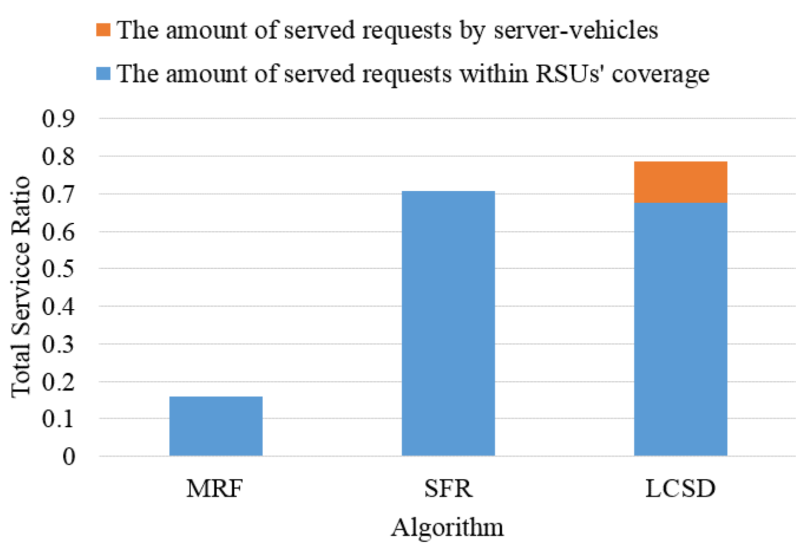


Figure 6. Total Service Ratio in Scenario I-80

Figures 5 and 6 show the total service ratio of the algorithms in scenarios US-101 and I-80, respectively. The scenarios have different statistics in terms of mean flow and speed. US-101 has a higher mean speed, so more vehicles pass the highway. X-axis represents the algorithms. Y-axis denotes the service ratio. To be specific, the service ratio is divided based on the service area: within RSUs' coverage (by the hybrid of I2V and V2V communications) and out of their coverage (by server-vehicles). As shown in the two figures, server-vehicles contribute to the considerable data service, so that *LCSD* outperforms both SFR and MRF in all scenarios. The overall total service ratio of the algorithms is higher in I-80 since the vehicles stay longer in the RSUs' coverage.

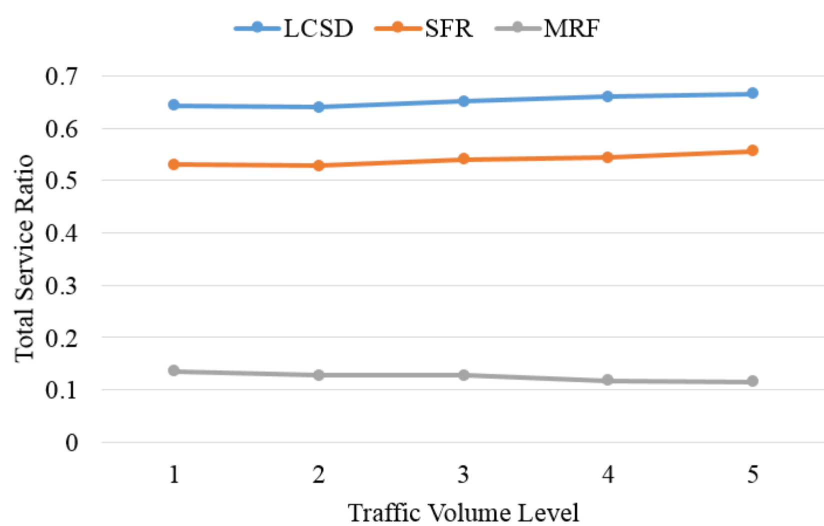


Figure 7. Total Service Ratio in Greenshield's model

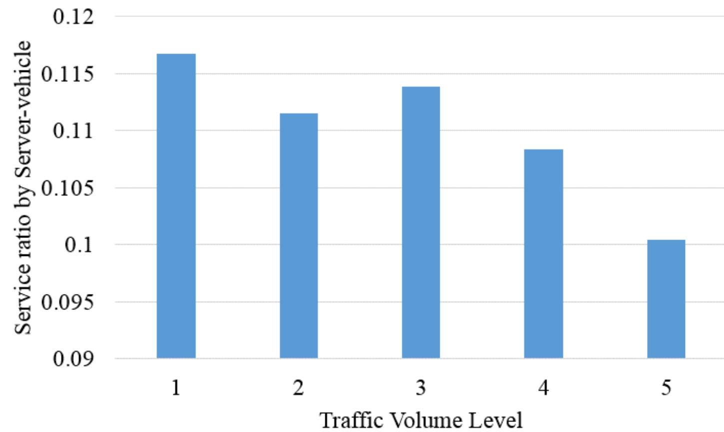


Figure 8. Service Ratio by Server-vehicle in Greenshield's model

Figure 7 shows the total service ratio of the algorithms in the Greenshield's model. X-axis represents the traffic volume level represented in TABLE 3. Figure 8 denotes the degree of the requests served by server-vehicles out of the total requests. Considering these figures, *LCSD* outperforms both SFR and MRF in all traffic volume levels by means of server-vehicles. Namely, *LCSD* has a better service ratio than SFR and MRF regardless the arrival rate or the average speed.

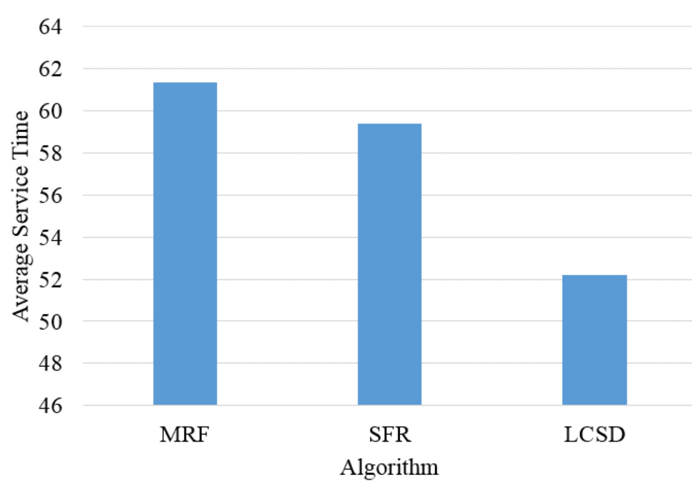


Figure 9. Average Service Time in Scenario US-101

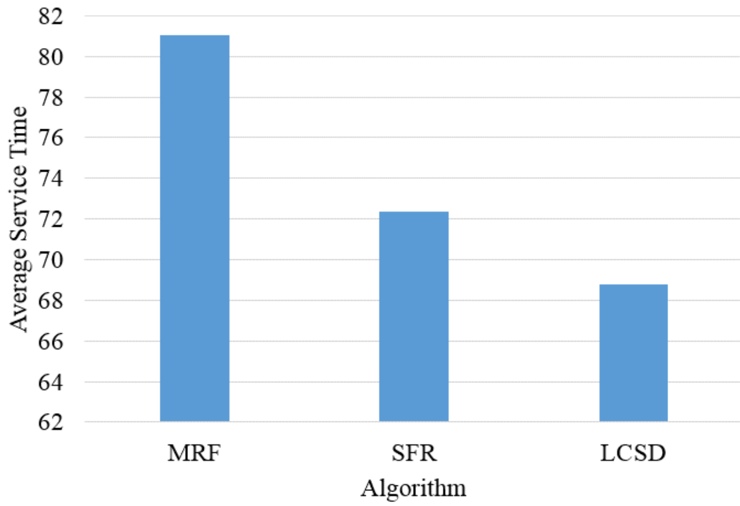


Figure 10. Average Service Time in Scenario I-80

Figure 9 and 10 show the average service time in scenarios US-101 and I-80, respectively. X-axis represents the algorithms. Overall, *LCSD* has less average service time than MRF and SFR, which means *LCSD* provides services more quickly than the other algorithms. Note that *LCSD* serves more requests than the other algorithms as shown in Figure 5 and 6. Therefore, it is non-trivial for *LCSD* to achieve shorter service time at the same time.

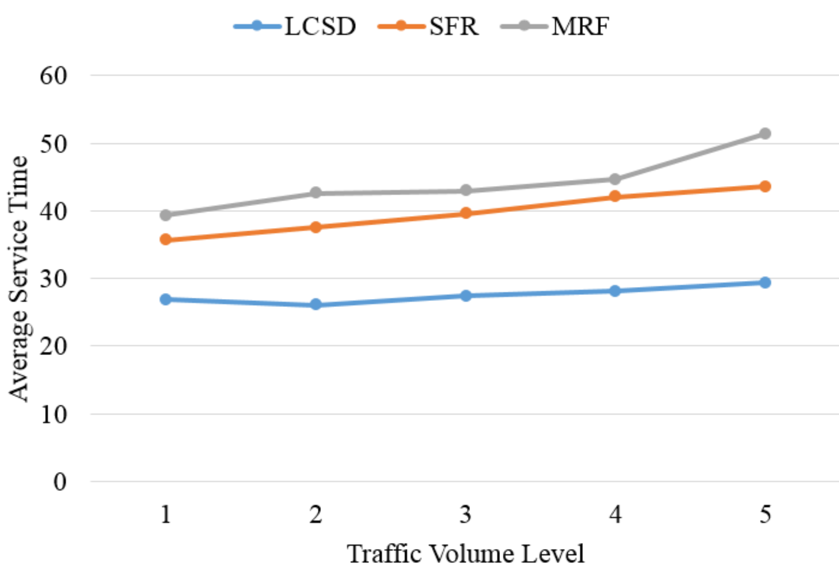


Figure 11. Average Service Time in Greenshield's model

Figure 11 represents the average service time in the Greenshield's model. X-axis represents the traffic volume level represented in TABLE 3. *LCSD* has a shorter average service time than MRF and SFR in every level.

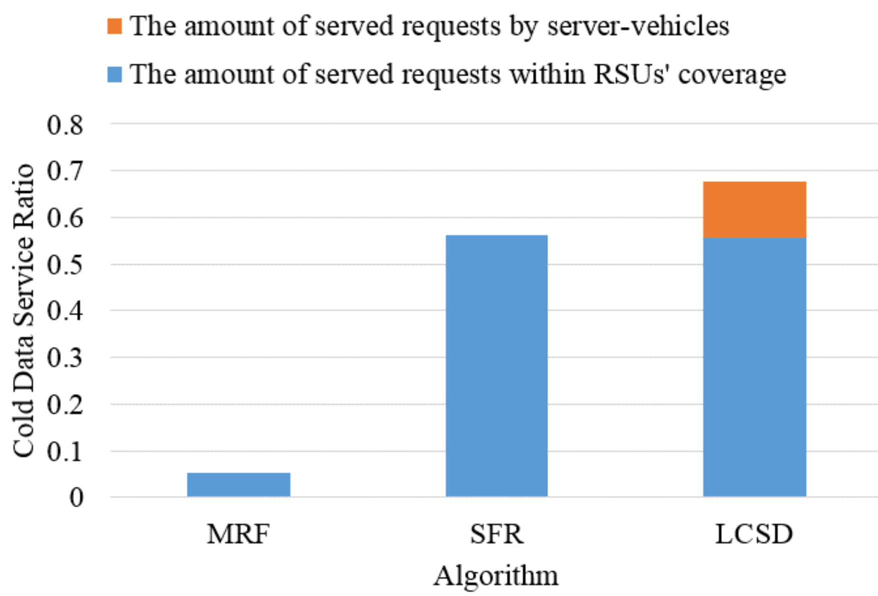


Figure 12. Cold Data Service Ratio in Scenario US-101

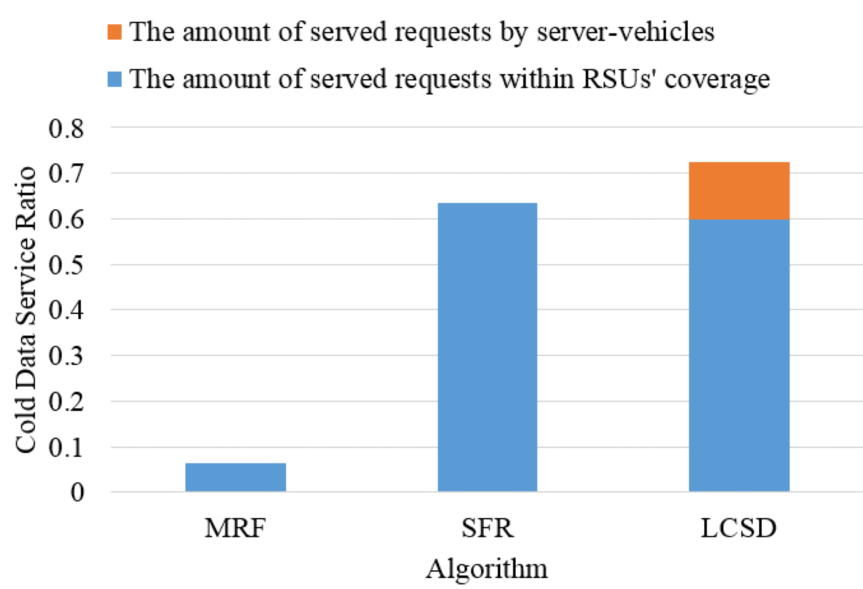


Figure 13. Cold Data Service Ratio in Scenario I-80

Figures 12 and 13 show the cold data service ratio of the algorithms in scenarios US-101 and I-80, respectively. X-axis represents the algorithms. Y-axis denotes the cold data service ratio, which is divided based on the service area: within RSUs' coverage (by the hybrid of I2V and V2V communications) and out of their coverage (by server-vehicles). As shown in the two figures, server-vehicles contribute to improve the service ratio, and *LCSD* has a better performance than MRF and SFR algorithms, regardless the scenarios.

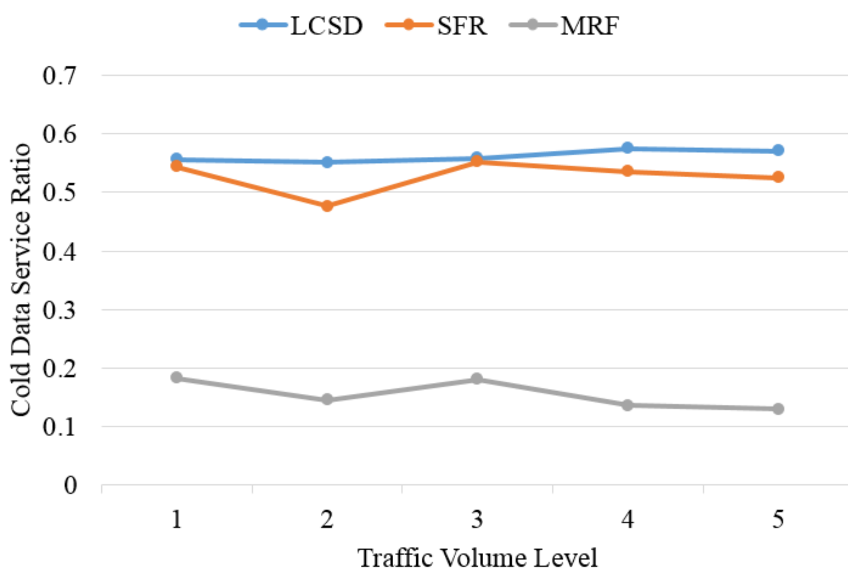


Figure 14. Cold Data Service Ratio in Greenshield's model

Figure 14 represents the cold data service ratio in the Greenshield's model. X-axis represents the traffic volume level represented in TABLE 3. Overall, LCSD has a slightly better result than MRF and SFR in every level. However, the difference of service ratios between LCSD and SFR is not as much as the difference value in Figures 12 and 13. This is because the number of lanes on the road is different.

Considering the above results, *LCSD* is scalable algorithms in terms of the total service ratio, the average service time, and the cold data service ratio, because the patterns of the performances are consistent in any traffic environment (e.g. US-101, I-80, and various traffic volume levels based on the Greenshield's model).

VI. Conclusion and Future work

In this work, we present the data dissemination system in vehicular networks based on the cooperation among multiple RSUs. Specifically, we propose a data dissemination solution via the hybrid of I2V and V2V communications in RSU's coverage and a server-vehicle designation strategy for the data service via V2V communication out of RSUs' coverage. On this basis, we further extend the solution to improve channel efficiency and offload workloads by transferring unserved requests to peer RSUs, so that the peer RSU, based on its own particular service status, may assign certain server-vehicles to assist the data service via V2V communication outside the RSU's coverage. Through a comprehensive simulation study, we verify that *LCSD* has the best performance in terms of enhancing the service ratio, improving the service fairness, and reducing the average service time.

In the future work, we will further investigate data dissemination problems in more sophisticated situations, including real-time services, heterogeneous databases and the dependency of multiple requested data items, etc. Meanwhile, the multi-hop V2V data dissemination is another critical issue to be considered for further enhance the system performance.

APPENDIX A

The formats of BSM and PVD messages are described in SAE J2735 message set dictionary. A given J2735 message is the payload of the next lower layer protocol, e.g. the “WAVE short message (WSM) data” field defined in IEEE 1609.3 [4]. Figure 15 shows the WSM packet format [27]. The maximum size of WSM packet is 1.4 Kbytes. In our system, we do not add additional information to the BSM. However, supplemental information, which is not defined in the message set dictionary such as the lists of the cached data items and the outstanding requests and information of neighboring vehicles, is added to PVD message. The WSM data size of PVD including the full optional contents is about 220 bytes according to [4]. About 280 IDs can be included in the PVD message, considering the vehicle ID size (4 bytes), transmission range (150m), vehicle size (5m), and safe distance (10m). It is enough number so as to include the list of neighboring vehicles in the PVD message.

The format of WSA is described in Figure 16. In our system, WSA contains diverse information such as scheduling result, an identifier of a server-vehicle, and a list of cached contents or outstanding request. Given that every vehicle cannot transmit or receive a data item within RSU’s coverage because of the characteristic of wireless communication, it is reasonable to add the additional information to WSA although many vehicles are within RSU’s coverage.

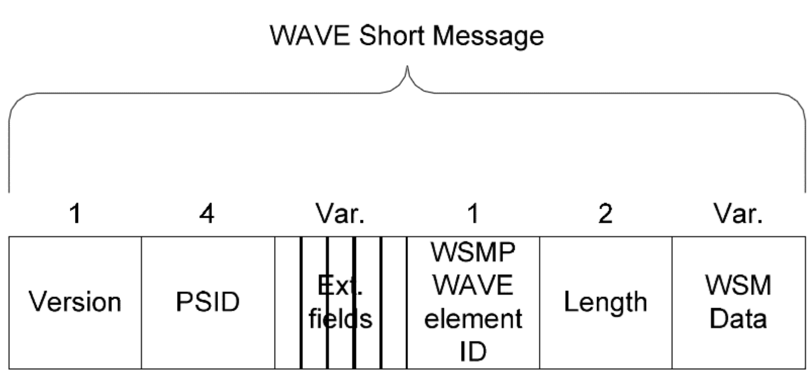


Figure 15. WSM packet format

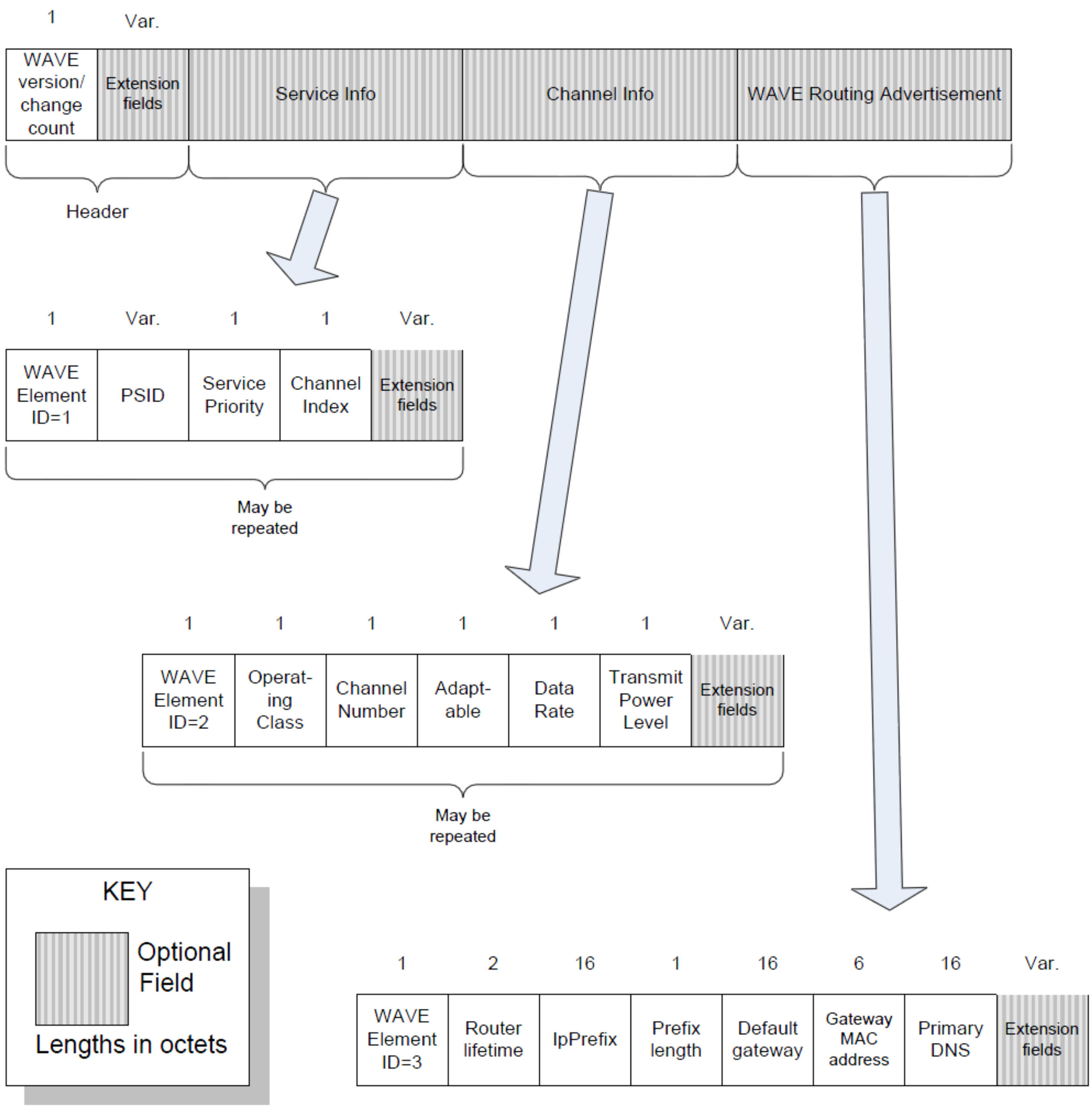


Figure 16. WSA packet format [27]

APPENDIX B

The computation time of *LCSD* exponentially increases along with the growing number of vertices or vehicles, which is represented in Figure 17 and 18. According to the figures, *LCSD* cannot finish the scheduling within 1 second when the number of vertices is more than about 1000. Given the computation time, the transmission ranges of RSU and OBU and the maximum number of requests generated within RSU's coverage need to be revised on the basis of the number of lanes on the road in order to finish the scheduling in time. Note that the computation time represented in Figure 17 and 18 is not accurate because each RSU and vehicle is controlled by each process on the simulation (e.g., Two RSUs implement *LCSD* simultaneously). Therefore, it is expected to take less time in the real environment.

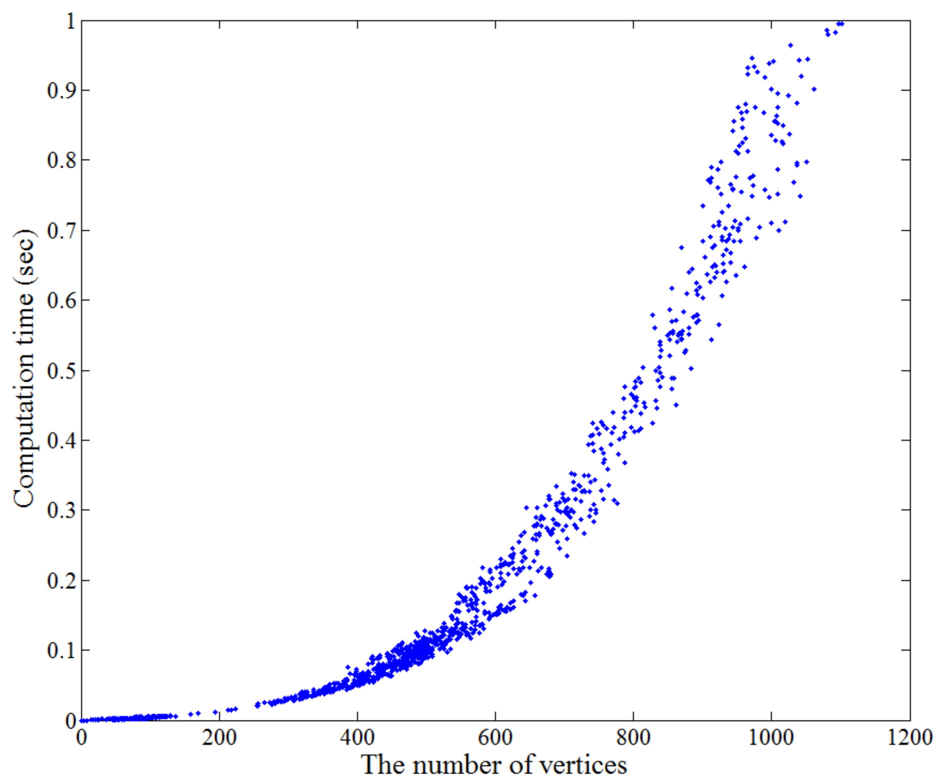


Figure 17. Computation time of *LCSD* along with the number of vertices

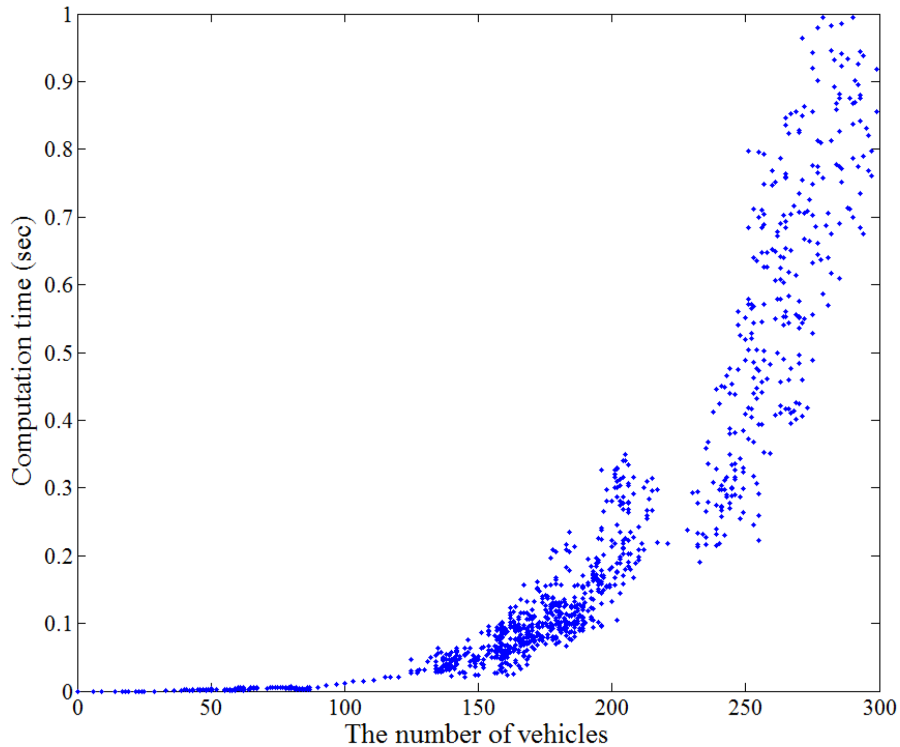


Figure 18. Computation time of *LCSD* along with the number of vehicles

REFERENCE

[1] K. Dar, M. Bakhouya, J. Gaber, M. Wack, and P. Lorenz, "Wireless Communication Technologies for ITS Applications," IEEE Communications Magazine, vol. 48, no. 5, pp. 156-162, May. 2010

[2] P. Papadimitratos, A. La Fortelle, K. Evenssen, R. Brignolo, S. Cosenza, and C. Ricerche, "Vehicular communication systems: Enabling technologies, applications, and future outlook on intelligent transportation" IEEE Communications Magazine, vol. 47, no. 11, pp. 84-95, Nov. 2009

[3] 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, 2010.

[4] *J2735 Dedicated Short Range Communications (DSRC) Message Set Dictionary*, SAE J2735, 2009.

[5] M. Zahren. (2013). *The Global Vehicle Population Aftermarket potential for remanufactured parts*. Available FTP: http://www.mera.org Directory: Document-Vault/Presentations/2013 File: Polk-Presentation-ReMaTec2013-FINAL.pdf

[6] K. Liu, J. Ng, V. Lee, W. Wu, S. H. Son, "Towards Scalable, Fair and Robust Data Dissemination via Cooperative Vehicular Communication," in *Embeded and Real-Time Computing Systems and Applications (RTCSA)*, Chongqing, 2014, pp. 1-9.

[7] Y. Li, X. Zhu, D. Jin, D. Wu, “Multiple Content Dissemination in Roadside Unit Aided Vehicular Opportunistic Networks”, IEEE Trans. on Vehicular Technology, vol. 63, no. 8, pp. 3947-3956, Oct. 2014.

[8] K. Liu, S. S. Lee, J. Ng, J. Chen, S. H. Son, “Temporal Data Dissemination in Vehicular Cyber-Physical Systems”, IEEE Trans. Intelligent Transportation Systems, vol. 15, no. 6, pp. 2419-2431, Dec. 2014.

[9] G. G. M. N. Ali and E. Chan, “Co-operative Data Access in Multiple Road Side Units (RSUs)-based Vehicular Ad Hoc Networks (VANETs),” in *Proc. Australasian Telecommunication Networks and Applications Conf.*, Melbourne, 2011, pp. 1-6.

[10] G. Ali, E. Chan, W. Li, “On scheduling Real-time Multi-item Queries in Multi-RSU Vehicular Ad Hoc Networks (VANETS)”, in *Advanced Information Networking and Applications Workshops (WAINA)*, Barcelona, 2013, pp.7-12.

[11] Y. Gui, E. Chan, “Data Scheduling for Multi-item Requests in Vehicle-Roadside Data Access with Motion Prediction Based Workload Transfer”, in *Advanced Information Networking and Applications Workshops (WAINA)*, Fukuoka, 2012, pp. 569-574.

[12] Y. Fallah, C.-L. Huang, R. Sengupta, and H. Krishnan, “Analysis of information dissemination in vehicular ad-hoc networks with application to cooperative vehicle safety systems,” IEEE Trans. Vehicular Technology, vol. 60, no. 1, pp. 233 –247, Jan. 2011.

[13] F.J. Ros, P.M. Ruiz, and I. Stojmenovic, “Acknowledgment-Based Broadcast Protocol for Reliable and Efficient Data Dissemination in Vehicular Ad-Hoc Networks,” IEEE Trans. Mobile Computing, vol. 11, no. 1, pp. 33-46, Jan. 2012.

[14] K. Fujimura and T. Hasegawa, “A Collaborative MAC Protocol for Intervehicle and Road to Vehicle Communications,” in *Proc. of the 7th International IEEE Conference on Intelligent Transportation Systems (ITSC’04)*, Washington D.C, 2004, pp. 816–821.

[15] C.-J. Chang, R.-G. Cheng, H.-T. Shih, and Y.-S. Chen, “Maximum freedom last scheduling algorithm for downlinks of DSRC networks,” IEEE Trans. Intell. Transp. Syst., vol. 8, no. 2, pp. 223–232, Jun. 2007.

[16] J. Wang, K. Liu, E. H. M. Sha, V. Lee, S. H. Son, "An Efficient Cluster-based Data Sharing Algorithm for Bidirectional Road Scenario in Vehicular Ad-hoc Networks,"*in Proceedings of the 15th International Conference on Algorithms and Architectures for Parallel Processing (ICA3PP 2015)*," Zhangjiajie, China, November 18-20, 2015.

[17] Q. Wang, P. Fan, and K. B. Letaief, “On the joint V2I and V2V scheduling for cooperative VANETS with network coding,” IEEE Trans. Veh. Technol., vol. 61, no. 1, pp. 62–73, Jan. 2012.

[18] G. Ali, E. Chan, and W. Li, “On scheduling data access with cooperative load balancing in vehicular ad hoc networks (VANETs),” J. Supercomput., vol. 67, no. 2, pp. 438–468, 2014.

[19] M. Krohn, R. Daher, M. Arndt, and D. Tavangarian, “Aspects of roadside backbone networks,” *in International Conf. on Wireless Communication, Vehicular Technology, Information Theory and Aerospace & Electronic Systems Technology*, Aalborg, 2009, pp. 788–792.

[20] H. Schwetman, "Csim19: a powerful tool for building system models," *in Proceedings of the 33nd Conference on Winter Simulation (WSC'01)*. IEEE, 2001, pp. 250–255.

[21] Systematics, Cambridge. NGSIM US-101 Data Analysis, Summary Reports, Federal HighWay Administration (FHWA). 2005.

[22] Systematics, Cambridge. NGSIM I-80 Data Analysis, Summary Reports. Federal High Way Administration (FHWA). 2005.

[23] K. Liu, E. Chan, V. Lee, K. Kapitanova, and S. H. Son, "Design and evaluation of token-based reservation for a roadway system," Transportation Research Part C: Emerging Technologies, vol. 26, pp. 184–202, Jan. 2013.

[24] IEEE Standard for Wireless Access in Vehicular Environments (WAVE)-Multi-channel Operation, IEEE Std 1609.4-2010, 2010.

[25] G. Zipf, Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology. Cambridge, Mass.: Addison-Wesley Press, 1949.

[26] J. W. Wong, "Broadcast delivery", Proceedings of the IEEE, vol. 76, no. 12, pp. 1566-1577, 1988.

[27] IEEE Standard for Wireless Access in Vehicular Environments (WAVE)-Networking Services, IEEE Std 1609.3-2010, 2010.

요약문

채널의 효율성 개선을 위한 데이터 보급 시스템

최근, 차량 애드혹 네트워크를 위한 통신의 발달은 운전을 위한 다양한 어플리케이션 구상에 많은 영향을 끼쳤을 뿐만 아니라 그러한 서비스들이 실현이 될 수 있도록 도움을 준다. 이때, 최근 구상 되는 어플리케이션들을 서비스 해주기 위해서는 어떻게 데이터 보급을 하느냐가 중요한 역할을 하게 된다. 하지만, 제한된 채널 환경과 차량들의 역동적인 움직임들로 인해 데이터를 차량들에게 효율적으로 보급 하는 것은 결코 간단한 일이 아니다. 이를 위해, 이 논문에서는 차량간 통신, 차량과 노변 장치간의 통신, 다중 노변 장치들간의 협업을 통한 효율적인 데이터 보급 시스템을 제안하고자 한다. 데이터 서비스를 최대한 보장을 해주기 위해서 두 가지의 목표를 가진다. 첫 번째는 차량간 통신과 차량과 노변 장치간 통신에 사용되는 채널을 효율적으로 사용하는 것이고, 두 번째는 노변장치에 부하되는 작업량을 줄이는 것이다. 이를 위해 우리는 다음과 같은 세 가지 접근법을 이용한다. 우선, 노변장치의 통신 반경 이내에서는 차량간 통신과 차량과 노변 장치간의 통신을 동시에 이용한 데이터 서비스 방법, 두 번째는 노변장치의 통신 반경이 미치지 않는 곳에서 데이터 서비스를 효율적으로 해주기 위해 서버 차량이라고 불리는 특정한 차량들만 데이터 서비스를 해주도록 허용해 주는 것, 세 번째는 이러한 차량을 지정하기 위해 노변 장치들이 서로 협업을 하는 것이다. 위의 세 가지 특성을 가진 시스템은 다른 알고리즘과 비교를 하였을 때 더욱 더 많은 양의 서비스를 제공해준다는 점과 서비스를 해주는데 평균적으로 걸리는 시간이 짧다는 점, 그리고 서비스에

대한 공정성이 더욱 좋다는 점을 통해서 우리가 제안하는 시스템의 우수성을 증명한다.

핵심어: 데이터 보급, 노변장치간 협업, 스케줄링, 차량간 통신과 차량과 노변장치간의 통신의 동시 사용