An SMDP-based Prioritized Channel Allocation Scheme in Cognitive Enabled Vehicular Ad Hoc Networks

Mushu Li, Lian Zhao, Hongbin Liang

***Abstract*—An efficient channel allocation scheme for the vehicular networks is required since the popularization and rapid growth of the corresponding applications in recent years. We propose a channel resource allocation scheme based on a semi-Markov decision policy (SMDP) to provide a solution for the problem of the channel resource shortage in the vehicular ad hoc networks (VANET). In this paper, we consider the channel allocation problem under a cognitive enabled vehicular ad hoc network environment. By a semi-Markov decision process, the channel allocation decision is made to maximize the overall system rewards. Besides, we consider services from two categories, primary users (PUs) and secondary users (SUs). On the top of the overall rewards maximization, we give priority to PU services without blocking any PU requests via cooperation between the roadside units (RSUs) and the base station. Numerical results and evaluations are presented to illustrate the desired performance of the proposed channel allocation scheme.**

I. Introduction

The vehicular ad hoc network (VANET) is envisioned to play a critical role in the intelligent transportation system (ITS) due to its safety, comfort in its traffic services. In ITS system, each vehicle attempts to exchange data with other vehicles and public facilities, which includes road safety messages and entertainment applications for passengers. VANET supports ITS system with high speed and wide range data transmission by vehicular communication. There are two major types of vehicular services in VANET: vehicle-to-vehicle (V2V) services and vehicle-to-infrastructure (V2I) services. In V2I communication, the vehicle is able to transmit data both downlink and uplink from roadside units (RSUs) with minimal latency [1]. Related data services include but not limited to hazard warnings, information on the current traffic situation, multimedia services and advertisements [2]. Compared with the traditional cellular networks, the vehicles on the road can enjoy the higher transmit rate in V2I services since RSUs are usually allocated along or close to the road. However, with the increasing number of various service requests, the bandwidth demand is growing up. The problem of spectrum limitation has gradually appeared. It is challenging for an RSU to allocate highly dynamic service requests with limited channels. Moreover, some vehicles in the system provide essential services related to public safety [3], [4], such as OBUs (vehicular On-Board Units) in vehicles of police, fire trucks, and ambulances. Compared with other regular vehicular users, they have the higher priority in the vehicular networks. It requires the RSU always provides enough channel resources to ensure those service accessibility to the spectrum.

Cognitive Radio (CR) is a context-aware intelligent radio, which represents a much broader paradigm in the wireless system since it enables improving the channel utilization and overall availability of the bandwidth [5]. Cognitive enabled vehicular network is proposed as a solution for overcoming the limitation of channel utilization in RSUs [6]. We can classify the service requests into primary users and secondary users in the vehicular networks. Primary users (PUs) are licensed users who are able to access the band with Quality-of-Service (QoS) guaranteed. In our work, the vehicular users who have the higher priority are regarded as PUs, such as OBUs in emergency service vehicle and mobile studio vehicles. Secondary users (SUs) sense the spectrum holes and find an opportunity to access the channel which PUs are not occupied [7]. Comparatively, users with lower priority services are secondary users. SUs can occupy the channel when all PUs' requests are satisfied, and the network provides QoS provision by the cooperation with the base station covering the RSU and vehicles. In our work, we propose a semi-Markov decision process (SMDP) policy to maximize channel utilization according to the long-term reward with improving QoS by the call admission control operation between requests from SUs and PUs in cognitive vehicular networks.

In recent years, there are many works studied the cognitive enabled vehicular networks. The authors in [8]–[10]

Mushu Li and Lian Zhao are with Department of Electrical and Computer Engineering, Ryerson University, Toronto, Canada ({mushu1.li,l5zhao@ryerson.ca}).

Hongbin Liang is with School of Transportation and Logistics, Southwest Jiaotong University, Chengdu, China (hbliang@swjtu.edu.cn).

proposed efficient spectrum-sensing methods for the cognitive vehicular network to minimize PU detection overhead. [11] investigated the cooperative spectrum sensing in a multi-channel CR network, which the channel condition and usage characteristics are considered. Furthermore, there are many related works proposed to improve road safety by acquiring extra bandwidth from those unused bands [9], [12]. In [9], the network architecture based on cognitive vehicular networking is designed to provide wireless connectivity to both the general public and emergency responders in emergency cases. Moreover, some works study spectrum resource management in cognitive vehicular network by different methods [6], [10], [11], [13]–[18], including game theory [15], Nash bargaining solution [10] and data mining [16]. In [11], the SU spectrum sharing problem is formulated to a weighted congestion game, and the relative algorithm is proposed to help SUs to achieve Nash Equilibrium. Auction mechanism is employed to solve the problem in [18]. The channel access is defined by a stochastic game considering the available channels and utility values of each vehicle. In [14], the problem of spectrum resource management on TV band white space in CR based high-speed vehicle network (CR-HSVN) is investigated. TV band white space in the vehicular networks also has been studied to solve route planning problem [19]. In [6], a semi-Markov decision process (SMDP) policy is proposed to improve video quality for video streaming services in the cognitive vehicular network. Video quality can be adjusted when the available spectrum is inadequate. In their work, PUs are fixed on the primary channel within the particular band as an ON and OFF model without participating in SMDP decision process, and the users can occupy primary channel only when it is available. Energy constraints in VANET also have been considered in [13].

In our work, we consider channel resource management in the heterogeneous cognitive vehicular network with controlled access between licensed users (PUs) request and unlicensed users (SUs) request by SMDP. We introduce an optimal channel allocation strategy to maximize the system reward in an RSU, in which both PUs and SUs are participating in the channel allocation. Moreover, we make a comparison between our model and the Greedy policy to present our improvement. The main contributions of this paper are listed as following:

1) The channel allocation problem is formulated to an SMDP model regarding a single RSU scenario in a cognitive vehicular network. In the system, the factors of user's income and service cost for occupying channel are considered in the reward strategy. For different request arrival rates, the system computes corresponding policies to maximize long-term rewards. Also, the decisions are able to be figured rapidly through historical records since the number of system states is finite. Therefore, it provides a highly dynamic solution for the problem.
2) Both PUs and SUs' QoS will be improved based on their arrival rate and channel availability. Channels are allocated conservatively when user arrival rate is high in order to admit more services. Otherwise, when user arrival rate is low, more channels are assigned to a request to improve the quality of the service. Thus, the system can offer higher transmit rate to a user if the channel resource is allowed.
3) In this paper, the spectrum in an RSU is entirely shared by PUs and SUs. Besides maximizing overall system reward, our SMDP model also provides a prioritized channel management policy to the services in the spectrum. We introduce the priority concept in SMDP model. PU requests are enabled to access the RSU spectrum anytime. SU requests and services are controlled by SMDP to guarantee the accessibility of PU requests. Compared to existing works, the proposed channel allocation strategy for PUs is more flexible. Therefore, in emergency cases, such as accident warning and nature disaster, RSUs always give priority to the specific users, such as emergency responders.

Comparing with existing literature in VANET, the main difference is following:

1) Comparing related literature [6], [20] which regards PU channel as the Markov decision process (MDP) model, we consider PU services arrival and departure as events due to the high mobility of PU, and PUs participate the channel allocation in SMDP policy with QoS guaranteed in our work. PU services are no longer fixed in the stationary primary channels of the spectrum. The system will maximize the overall reward both from PUs and SUs by our proposed policy.
2) Our policy prioritizes PU service requests in the decision process. When a PU request arrives, and the RSU spectrum is full, the RSU will handover one or more SU services to the base station, which covering the RSU with the broader spectrum to release channels for serving PU request. The corresponding action is made by our SMDP model aiming to achieve overall system reward maximization without sacrificing PU services or the interference from SUs by our proposed policy, which can improve the spectrum utilization significantly. There is no priority selection in existing works who studied resource allocation by the SMDP model [6], [20]–[22].

The remaining of the paper is organized as follows. In Section II, the system model of the cognitive vehicular network is introduced. The channel allocation model based on SMDP for a single roadside unit is discussed in Section III. In Section IV, we present the proposed algorithm to implement our model. Then, the system performance is evaluated in Section V. Finally, conclusions and the future works are given in Section VI.

II. System Model

A. Network Structure

Fig. 1 shows the network structure in a one-way highway scenario. There are N_R cognitive enabled RSUs employed along the road, and each RSU is covering a part of the road with d_R coverage diameter. All N_R RSUs and their users under a base station coverage, such that user can always transmit and receive data either from the adjacent RSU or the base station. We consider that the residence time for a vehicle associated with an RSU follows an exponential distribution with a mean time of $1/\mu_d$ [6], [20].

Considering two categories of users: i) Primary Users: Primary vehicle users have priority to transmit in the licensed spectrum bands, such as emergency command vehicles and mobile studio vehicles. ii) Secondary Users: Unlicensed vehicle users who sense the available channels and attempt to use primary band is called secondary users. In our case, SUs and PUs share the licensed channels, denoted as C_n, $n \in \{1, \ldots, K\}$. PUs have the priority to access the spectrum, while SUs can occupy the empty channels when those channels are unoccupied by PUs. We assume that both types of users arrive with Poisson distribution, with the mean rate of λ_p for PUs and λ_s for SUs. Service time follows an exponential distribution. With one allocated channel use, average service time is $1/\mu_p$ for PU and $1/\mu_s$ for SU. For example, if a PU service is allocated to c channels, the mean service time is $1/(c\mu_p)$.

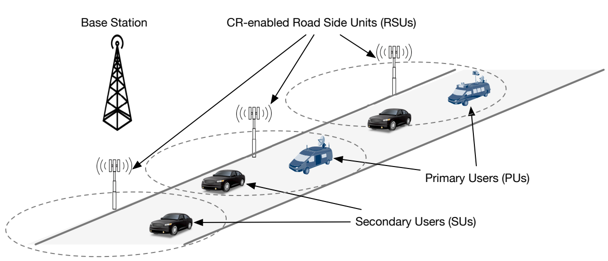


Fig. 1. System model for cognitive vehicular network.

B. Channel Model

Once a user under an RSU coverage plans to use channels of the RSU, the RSU will detect the type of request and decide whether to accept it or reject it based on channel resources availability. If the request is accepted, the RSU allocates the available channels to the user. A user can obtain the higher transmission rate if more channels are assigned to the service, then the user has less cost for occupying the channels since its services can be completed in a shorter time period. While, if the user occupies the full spectrum, more services will be rejected and user's satisfaction will decrease correspondingly. In our model, we balance the user satisfaction and service cost simultaneously. The system will distribute available channels to new request first, no matter the request from SU or PU. Otherwise, if none of the channels is available, the RSU will search the channels which are allocated to SUs, and the system will stop the appropriate SU services in order to offer the channel space to the PU. The affected SU's request will be transferred to the base station. In the remaining of the paper, the action of transfer indicates that the SU services are handoff from an RSU to base station passively by the RSU for reserving channel resource to PU services. The event of handoff means that the services are handoff from an RSU to base station or another RSU since the SUs are out of the RSU coverage.

Assume that a roadside unit has K channels. The number of channels which can be allocated to one service is c, where $c \in \{1, \ldots, C\}$, $C \leq K$. C is the maximum number of the channels allocated to one service. We assume that one channel can meet minimum service requirement of all kinds of services. To ensure the system accessible and stable for primary users, $\lambda_p < K\mu_p$. When all of the channels are busy, since the system always provides the channel resources to primary users, we may need to transfer SU's service which occupying the RSU channels to the base station or shrink channels of another PU service. Fig. 2 shows an example of channel allocation when channels are busy. When a request from a primary user arrived in (a), one or more services for SUs determined by SMDP policy will be transferred to the base station with the certain cost to accept the priority request, which shows in (b). Then, in some extreme scenarios, when all channels are occupied by primary users and new primary user arrives, the system will find the service occupying two or more channels and shrink it to accept a new request, which shows in (c). We consider the PU channel degradation as an operation to ensure the channels accessibility for all PUs for the extreme case in the stochastic process, and it brings no cost to the system. All the services which need to be transferred are determined by the proposed SMDP policy which will be introduced in next section to maximize the long-term system reward.

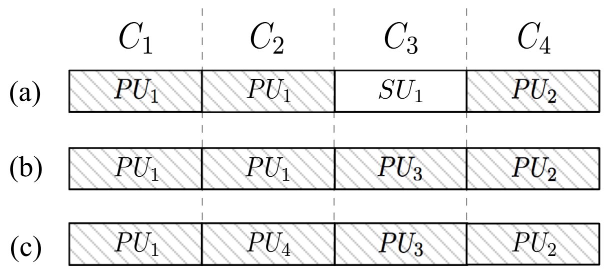


Fig. 2. An Example of channel allocation when channels are busy.

III. SMDP Formulation for a Single RSU

In this section, we propose an SMDP-policy to optimize the channel allocation regarding the system reward in a dynamic network with a single RSU. There are five parts in an SMDP model: a) system states and events; b) set of actions; c) decision epochs; d) transition probabilities and e) reward model.

A. System States

We consider two types of services participating the channel allocation, PUs and SUs. The channels allocated to the PU services is denoted by a vector $\boldsymbol{n}_p = [n_{p1}, \ldots, n_{pC}]^T$, where n_{pc} represents the number of PU services with c channels, and $\sum_{c=1}^{C} cn_{pc} \leq K$. Similarly, the number of channels allocated to the SU services is denoted by another vector $\boldsymbol{n}_s = [n_{s1}, \ldots, n_{sC}]^T$. The events e are summarized in Table I. Therefore the system state can be defined as:

$$S_v = \{s_v | s_v = \langle \boldsymbol{n}_s, \boldsymbol{n}_p, e \rangle\} \tag{1}$$

where $\sum_{c=1}^{C}[c(n_{pc} + n_{sc})] \leq K$.

TABLE I
Summary of Events

e	Event
A_p	PU service arrival
A_s	SU service arrival
F_{pc}	PU service with c channels completed or hand-off
F_{sc}	SU service with c channels completed or hand-off

B. Set of Actions

After the system receives a new request, the system will determine the following actions. If the request is from a SU and it has been accepted with c channels, it denotes as $a(s_v) = (s, c)$. Otherwise, $a(s_v) = (p, c, \boldsymbol{T})$ indicates that the request is from a PU and accepted with c channels, where $\boldsymbol{T}$ is transfer vector for SUs. T_c represents the number of SU services in c channel is transferred, where $\sum_{c=1}^{C} cT_c \leq K$, $c \in \{1, \ldots, C\}$. If the request is rejected, the action is $a(s_v) = 0$. When the service is completed or leaving the RSU's coverage area, the channels will be released, and such action is denoted as $a(s_v) = -1$. The action space $\mathbb{A}(s_v)$ is summarized as follows:

$$\mathbb{A}_(s_v) = \begin{cases} -1, & e \in \{F_{sc}, F_{pc}\} \\ \{0, (s, c)\}, & e = A_s \\ (p, c, \boldsymbol{T}). & e = A_p \end{cases} \tag{2}$$

C. Decision Epochs

The epoch of the transition to next state j_v, when an action $a(s_v)$ is selected in current state s_v, is analysed in this section. Since the service arrival follows Poisson distribution, we first determine the mean rate $\gamma(s_v, a)$ of events, which is expressed as Eq. (3). Then, the expected time interval between state s_v and j_v is given by $\frac{1}{\gamma(s_v,a)}$, which follows an exponential distribution.

In Eq. (3), λ_p and λ_s are the arrival rate for PUs requests and SUs requests respectively. When the system doesn't accept any request, there are $\sum_{c=1}^{C}(n_{sc} + n_{pc})$ existing services in the system, then the departure rate of the overall system is $\sum_{c=1}^{C}(cn_{sc}\mu_s + cn_{pc}\mu_p)$, and the hand-off rate is $\sum_{c=1}^{C}(n_{sc} + n_{pc})\mu_d$. If a SU request is admitted by c channels, then there are $\sum_{c=1}^{C}(n_{sc} + n_{pc}) + 1$ services in the system. The new departure and handoff rate are $\sum_{c=1}^{C}(cn_{sc}\mu_s + cn_{pc}\mu_p) + c\mu_s$ and $\sum_{c=1}^{C}(n_{sc} + n_{pc})\mu_d + \mu_d$ respectively. Similarly, if a PU request is admitted by c channels with $\boldsymbol{T}$ transferred, then the system has $\sum_{c=1}^{C}[(n_{sc} - T_c) + n_{pc}] + 1$ services. The rate of departure is $\sum_{c=1}^{C}[c(n_{sc} - T_c)\mu_s + cn_{pc}\mu_p] + c\mu_p$, and the rate of handoff is $\sum_{c=1}^{C}(n_{sc} - T_c + n_{pc})\mu_d + \mu_d$.

D. Transition Probabilities

After we determined the mean rate of the events, $q(j_v|s_v, a)$ can be defined as the transition probability from state s_v to next state j_v when action a is occurred. We have a few cases of the $q(j_v|s_v, a)$ depending on the states of s_v as following.

When $s_v = \langle\ \boldsymbol{n}_s,\ \boldsymbol{n}_p,\ A_p\ \rangle$,

$$q(j_v|s_v, a) = \begin{cases} \frac{(n_{pc}+1)(c\mu_p+\mu_d)}{\gamma(s_v,a)}, & j_v = \langle \boldsymbol{n}_s - \boldsymbol{T}, \boldsymbol{n}_p, F_{pc} \rangle \\ \frac{n_{pm}(m\mu_p+\mu_d)}{\gamma(s_v,a)}, & j_v = \langle \boldsymbol{n}_s - \boldsymbol{T}, \boldsymbol{n}_p + \boldsymbol{I}_c - \boldsymbol{I}_m, F_{pm} \rangle \\ \frac{n_{sm}(m\mu_s+\mu_d)}{\gamma(s_v,a)}, & j_v = \langle \boldsymbol{n}_s - \boldsymbol{T} - \boldsymbol{I}_m, \boldsymbol{n}_p + \boldsymbol{I}_c, F_{sm} \rangle \\ \frac{\lambda_s}{\gamma(s_v,a)}, & j_v = \langle \boldsymbol{n}_s - \boldsymbol{T}, \boldsymbol{n}_p + \boldsymbol{I}_c, A_s \rangle \\ \frac{\lambda_p}{\gamma(s_v,a)}, & j_v = \langle \boldsymbol{n}_s - \boldsymbol{T}, \boldsymbol{n}_p + \boldsymbol{I}_c, A_p \rangle \end{cases} \tag{4}$$

$$\gamma(s_v,a)=\tau(s_v,a)^{-1}=$$
$$\begin{cases}\lambda_p+\lambda_s+\sum\limits_{c=1}^{C}[n_{sc}(c\mu_s+\mu_d)+n_{pc}(c\mu_p+\mu_d)], & e\subseteq\{F_{sc},F_{pc}\}\text{ or}\\ & e\subseteq\{A_p,A_f\},\ a=0\\ \lambda_p+\lambda_s+\sum\limits_{c=1}^{C}[n_{sc}(c\mu_s+\mu_d)+n_{pc}(c\mu_p+\mu_d)]+(c\mu_s+\mu_d), & e=\{A_s\},a=(s,c),\\ \lambda_p+\lambda_s+\sum\limits_{c=1}^{C}[(n_{sc}-T_c)(c\mu_s+\mu_d)+n_{pc}(c\mu_p+\mu_d)]+(c\mu_p+\mu_d), & e=\{A_p\},a=(p,c,\boldsymbol{T}).\end{cases}\tag{3}$$

where $a=(p,c,\boldsymbol{T})$, $c\in\{1,\dots,C\}$, $m\in\{1,\dots,C\}$, and $c\neq m$. The vector $\boldsymbol{I}_c$ indicates a vector with C elements, where the c-th element is 1, others are 0. Similarly, in vector $\boldsymbol{I}_m$, the m-th element is 1, others are 0.

When state $s_v=\langle\ \boldsymbol{n}_s,\boldsymbol{n}_p,\ A_s\ \rangle$,

$$q(j_v|s_v,a)=$$
$$\begin{cases}\frac{\lambda_s}{\gamma(s_v,a)}, & j_v=\langle\boldsymbol{n}_s,\boldsymbol{n}_p,A_s\rangle,a=0\\ \frac{\lambda_p}{\gamma(s_v,a)}, & j_v=\langle\boldsymbol{n}_s,\boldsymbol{n}_p,A_p\rangle,a=0\\ \frac{n_{sc}(c\mu_s+\mu_d)}{\gamma(s_v,a)}, & j_v=\langle\boldsymbol{n}_s-\boldsymbol{I}_c,\boldsymbol{n}_p,F_{sc}\rangle,a=0\\ \frac{n_{pc}(c\mu_p+\mu_d)}{\gamma(s_v,a)}, & j_v=\langle\boldsymbol{n}_s,\boldsymbol{n}_p-\boldsymbol{I}_c,F_{pc}\rangle,a=0\\ \frac{(n_{sc}+1)(c\mu_s+\mu_d)}{\gamma(s_v,a)}, & j_v=\langle\boldsymbol{n}_s,\boldsymbol{n}_p,F_{sc}\rangle,a=(s,c)\\ \frac{n_{sm}(m\mu_s+\mu_d)}{\gamma(s_v,a)}, & j_v=\langle\boldsymbol{n}_s+\boldsymbol{I}_c-\boldsymbol{I}_m,\boldsymbol{n}_p,F_{sm}\rangle,\\ & a=(s,c)\\ \frac{n_{pm}(m\mu_p+\mu_d)}{\gamma(s_v,a)}, & j_v=\langle\boldsymbol{n}_s+\boldsymbol{I}_c,\boldsymbol{n}_p-\boldsymbol{I}_m,F_{pm}\rangle,\\ & a=(s,c)\\ \frac{\lambda_s}{\gamma(s_v,a)}, & j_v=\langle\boldsymbol{n}_s+\boldsymbol{I}_c,\boldsymbol{n}_p,A_s\rangle,a=(s,c)\\ \frac{\lambda_p}{\gamma(s_v,a)}, & j_v=\langle\boldsymbol{n}_s,\boldsymbol{n}_p+\boldsymbol{I}_c,A_p\rangle,a=(s,c)\end{cases}\tag{5}$$

where $c\in\{1,\dots,C\}$, $m\in\{1,\dots,C\}$, and $c\neq m$.

When the state $s_v=\langle\ \boldsymbol{n}_s,\boldsymbol{n}_p,\ e\ \rangle$, where $e\in\{F_{sc},F_{pc}\}$. There is no special action except releasing channels, where $a(s_v)=-1$. The corresponding transition probabilities can be expressed as

$$q(j_v|s_v,a)=\begin{cases}\frac{\lambda_s}{\gamma(s_v,a)}, & j_v=\langle\boldsymbol{n}_s,\boldsymbol{n}_p,A_s\rangle\\ \frac{\lambda_p}{\gamma(s_v,a)}, & j_v=\langle\boldsymbol{n}_s,\boldsymbol{n}_p,A_p\rangle\\ \frac{n_{sc}(c\mu_s+\mu_d)}{\gamma(s_v,a)}, & j_v=\langle\boldsymbol{n}_s-\boldsymbol{I}_c,\boldsymbol{n}_p,F_{sc}\rangle\\ \frac{n_{pc}(c\mu_p+\mu_d)}{\gamma(s_v,a)}. & j_v=\langle\boldsymbol{n}_s,\boldsymbol{n}_p-\boldsymbol{I}_c,F_{pc}\rangle\end{cases}\tag{6}$$

E. Reward Model

The system will be rewarded in system states and the corresponding actions. The reward function is formulated by the income from users $w(s_v,a)$ and the system cost $g(s_v,a)$, which occurred on action a in state s_v. Therefore, the reward $r(s_v,a)$ can be formulated as following:

$$r(s_v,a)=w(s_v,a)-g(s_v,a),\tag{7}$$

where $w(s_v,a)$ is evaluated as below,

$$w(s_v,a)=\begin{cases}0, & a(s_v)=-1,\\ & e\in\{F_{sc},F_{pc}\}\\ -\gamma_sU_s, & a(s_v)=0,\\ & e=A_s\\ \gamma_sU_s-\frac{\theta\beta}{c}, & a(s_v)=(s,c),\\ & e=A_s\\ \gamma_pU_p-\frac{\theta\beta}{c}-\sum\limits_{c=1}^{C}T_cE_t & a(s_v)=(p,c,\\ \qquad-\sum\limits_{c=1}^{C}cT_cU_t. & \boldsymbol{T}),e=A_p\end{cases}\tag{8}$$

In (8), there is no income from users when the channels are released, where $a(s_v)=-1$ and $e\in\{F_{sc},F_{pc}\}$. U_s and U_p are the satisfaction income from SU and PU respectively. γ_s and γ_p represent weight factors. When service arrives and system rejects the SU service ($a(s_v)=0$, $e=A_s$), the system will lose the satisfaction income from the user. Otherwise, the system will gain the user satisfaction when system accepts the service. Also, the transmission cost for occupying channels from users is taken into account when system accepts service. θ denotes the service transmission time, and β is the price per unit time for transmission on one channel. The term $\frac{\theta\beta}{c}$ indicates the user transmission cost for occupying c channels. Besides, if the PU's request is accepted when SUs' services are transferred, the cost of transferring service occurs. E_t denotes the transfer cost for one SU service, and U_t is the dropping cost for one channel. Thus $\sum\limits_{c=1}^{C}T_cE_t+\sum\limits_{c=1}^{C}cT_cU_t$ is the overall cost for transferring SU services.

The system cost $g(s_v,a)$ is

$$g(s_v,a)=o(s_v,a)\tau(s_v,a),\tag{9}$$

where $\tau(s_v,a)$ is the expected service time given by Eq. (3), and $o(s_v,a)$ is the cost rate of service time when an action $a(s_v)$ is selected. $o(s_v,a)$ can be determined by the number of occupied channels, shown as,

$$o(s_v,a)=\sum_{c=1}^{C} c(n_{sc}+n_{pc}). \tag{10}$$

Moreover, we can evaluate the expected discounted reward $r(s_v,a)$ based on the discounted reward model given by [23] as following,

$$\begin{aligned} r(s_v,a) &= w(s_v,a)-o(s_v,a)E_s^a\left\{\int_0^{\tau} e^{-\alpha t}dt\right\} \\ &= w(s_v,a)-o(s_v,a)E_s^a\left\{\frac{[1-e^{-\alpha\tau}]}{\alpha}\right\} \\ &= w(s_v,a)-\frac{o(s_v,a)}{\alpha+\gamma(s_v,a)}, \end{aligned} \tag{11}$$

where α is a continuous-time discounted factor.

From above transition probabilities and Eq. (11), the maximum long-term discounted reward can be formulated to Bellman equation with the discount reward model defined in [23] as below,

$$\nu(s_v)=\max_{a\in Act_s}\left\{r(s_v,a)+\lambda\sum_{j_v\in S} q(j_v|s_v,a)\nu(j_v)\right\}. \tag{12}$$

where $\lambda=\frac{\gamma(s_v,a)}{\alpha+\gamma(s_v,a)}$.

Furthermore, to achieve the unified expected system reward, we introduce a new parameter $\omega=\lambda_p+\lambda_s+KC(\mu_p+\mu_s)+\sum_{c=1}^{C}\lfloor\frac{K}{c}\rfloor\mu_d$. Then the unified transition probability can be formulated as following:

$$\bar{q}(j_v|s_v,a)=\begin{cases} 1-\frac{\gamma(s_v,a)[1-q(j_v|s_v,a)]}{\omega}, & j_v=s_v \\ \frac{\gamma(s_v,a)q(j_v|s_v,a)}{\omega}. & j_v\neq s_v \end{cases} \tag{13}$$

The reward function after uniformization is

$$\bar{r}(s_v,a)=r(s_v,a)\frac{\gamma(s_v,a)+\alpha}{\omega+\alpha}. \tag{14}$$

Then, according to (13) and (14), the maximization discount reward model problem can be expressed as

$$\bar{\nu}(s_v)=\max_{a\in Act_s}\left\{\bar{r}(s_v,a)+\bar{\lambda}\sum_{j_v\in S}\bar{q}(j_v|s_v,a)\bar{\nu}(j_v)\right\}, \tag{15}$$

where the uniformization parameter is

$$\bar{\lambda}=\frac{\omega}{\omega+\alpha}. \tag{16}$$

IV. Channel Allocation Scheme by SMDP

In this section, we propose our algorithm to solve Eq. (15). Firstly, we need to search all the possible actions Act_s for the finite state spaces s_v. The algorithm description is stated in Algorithm 1.

Algorithm 1 Searching available action for SMDP

```
Initialize action sets Act_s, Act_1, Act_2, Act_3 = ∅.
Set X ← {s_v | e(s_v) ∈ {F_sc, F_pc} }, Y ← {s_v | e(s_v) = A_sc}, Z ← {s_v | e(s_v) = A_pc}, c = 1, ..., C.
    Ensure that Σ_{c=1}^{C} c(n_sc + n_pc) ≤ K for state s_v.
if s_v ∈ X then
    Act_1 ← {a | a = −1}.
else if s_v ∈ Y then
    for m = 0 : C do
        Ω_s ← {s_v | Σ_{c=1}^{C} c(n_sc + n_pc) + m ≤ K }.
        if s_v ∈ Ω_s and m ≠ 0 then
            Act_2 ← {a | a = ⟨s, m⟩} ∪ Act_2.
        else
            Act_2 ← {a | a = 0} ∪ Act_2.
        end if
    end for
else if s_v ∈ Z then
    for m = 1 : C do
        Ω_p ← {s_v | Σ_{c=1}^{C} c(n_sc + n_pc) + m ≤ K }.
        if s_v ∈ Ω_p then
            Act_3 ← {a | a ∈ ⟨p, m, 0⟩} ∪ Act_3.
        else
            Find all possible T that
                Σ_{c=1}^{C} c(n_sc − T_c + n_pc) + m ≤ K,
                where T_c ≤ n_sc, c = 1, ..., C.
            Act_3 ← {a | a ∈ ⟨p, m, T⟩} ∪ Act_3.
        end if
    end for
end if
Return Act_s = Act_1 ∪ Act_2 ∪ Act_3, Act_s is the possible action for state s_v.
```

In Algorithm 1, Act_1 is the set of all actions releasing channels. Act_2 is the set of actions for accepting or rejecting the SU request. Act_3 is the set of actions for accepting PU request and transferring SU requests to the base station. After we find all possible actions for all state space s_v, then we solve Eq. (15) by the iteration reward value method. The detailed description is in Algorithm 2. The output d_{opt} is the decision policy of the system.

V. Performance Evaluation

In this section, the performance of the proposed SMDP policy is evaluated. The results are compared with Greedy policy [24] to show the enhancement of our method. In

Algorithm 2 Channel Allocation Scheme

1: Set long term reward $v(s_v) = 0$ of all states. And iteration $i = 0$.
2: For each s_v, find corresponding reward by Eq. (12). Act_s is obtained by Algorithm 1.
$$\bar{\nu}^{i+1}(s_v) = \max_{a \in Act_s} \left\{ \bar{r}(s_v, a) + \bar{\lambda} \sum_{j_v \in S} \bar{q}(j_v | s_v, a) \bar{\nu}^i(j_v) \right\}.$$
3: **if** $|\bar{\nu}^{i+1} - \bar{\nu}^i| > \frac{\varepsilon(1-\bar{\lambda})}{2\bar{\lambda}}$ **then**
4: Back to step 2, $i = i + 1$.
5: **else**
6: Find corresponding action policy for $\nu^{i+1}(s_v)$,
$$d_{opt} \in \arg \max_{a \in Act_s} \left\{ \bar{r}(s_v, a) + \bar{\lambda} \sum_{j_v \in S} \bar{q}(j_v | s_v, a) \bar{\nu}^i(j_v) \right\}.$$
7: **end if**

Greedy policy, the services are always allocated to the maximum available channels.

Suppose a cognitive enabled RSU system has K channels in the network. The maximum number of channels for one service is $C = 2$. The average departure rate of the SU service is $\mu_s = 3$, and the average departure rate of the PU service is $\mu_p = 2$. We suppose the handoff rate is much lower than the departure rate for both PU and SU. The setting of rewards is shown in Table II. The discount factor α is 0.1. The simulation is run by 100 sec, and it is repeated by 10 times to obtain the average performance value. SUs and PUs arrive in Poisson distribution with dynamic rate λ_s and λ_p. To make the simulation model reasonable and stable, we set $\lambda_s < K\mu_s$ and $\lambda_p < K\mu_p$.

TABLE II
SIMULATION PARAMETERS

U_s	U_p	E_t	U_t	γ_s	γ_p	θ	β
30	40	5	4	1	1	8	1

Figs. 3 to 5 are the simulation results of the action probabilities with different variable setting for two categories of users. The comparison results between our model and the Greedy policy are shown in Figs. 6 to 7. The sub-figures (a) for Figs. 3 to 7 are the simulation results when PU arrival rate λ_p increases from 1 to 10, $\lambda_s = 5$, $\mu_s = 3$, $\mu_p = 2$, and $K = 6$. The sub-figures (b) for Figs. 3 to 7 show the results when SU arrival rate λ_s is variable from 1 to 10, $\lambda_p = 2$, $\mu_s = 3$, $\mu_p = 2$, and $K = 6$, and (c) are the results when channel number K is changing from 2 to 11, $\lambda_p = 2$, $\lambda_s = 5$, $\mu_s = 3$, and $\mu_p = 2$.

The results of the action probabilities among all SU services are shown in Fig. 3. Fig. 3(a) shows the action probability distribution as a function of PU arrival rate λ_p. We can observe that most of the SUs are allocated to 2 channels for higher transmit rate when λ_p is low. When λ_p increases, more SU requests are allocated to 1 channel or blocked since it reserves channel resource to serve increasing arrival requests to gain a higher reward. A similar trend of action probabilities is observed in Fig. 3(b), when SU arrival rate λ_s is changing as a variable. Fig. 3(c) shows the influence for SU action probabilities from the channel number K. We see the more SU requests are allocated to fewer channels or blocked by SMDP policy when the channel number is low. The blocking rate is reducing, and the action probabilities for service allocated to more channel raises correspondingly when the channel number K increases. This is because that the policy considers three main factors to maximize system reward: the user satisfaction, the cost for users transmission and the cost of occupying channels. Our policy prefers to allocate more channels in order to reduce the service departure time and the related cost in low service arrival rate or sufficient channel resources. Otherwise, it allocates channels conservatively with the purpose of accepting more services when the service request rate is high, or channel resources are in shortage.

The results of the action probabilities among all PU services are presented in Fig. 4. Similar to Fig. 3, all or most of PU services are allocated to 2 channels in low request arrival rate without channel resource limitation. And when channel resources become strained, or arrival rate increases, our policy would like to perform the action of allocating service in fewer channels to accommodate more service requests. However, different with the results in SU, none of PU requests is rejected since our policy gives priority to PU service requests.

Fig. 5 depicts the transfer action of the SU services when the RSU channels are fully occupied. When no channel is available for accepting new PU requests, the RSU will transfer the corresponding SU services to the base station in order to provide the channel to PUs. As the arrival rate is increasing, more SU services have to be transferred to the base station covering the RSU and the vehicles, which is shown in Figs. 5(a) and (b). We see the probability of transferring service with 1 channel is increased greatly than the probability of transferring services with 2 channels since the cost of transferring services with fewer channels is lower than the cost of transferring services with multiple channels. Also, the number of services allocated in fewer channels is dominated when arrival rate is high, which is another reason for the behavior. On the other hand, if the channel number is increasing as shown in Fig. 5(c), fewer of SU services are transferred since the RSU is able to admit more services from both PUs and SUs simultaneously.

Fig. 6 and Fig. 7 present the comparison of our policy and the Greedy policy. The simulation results for unified

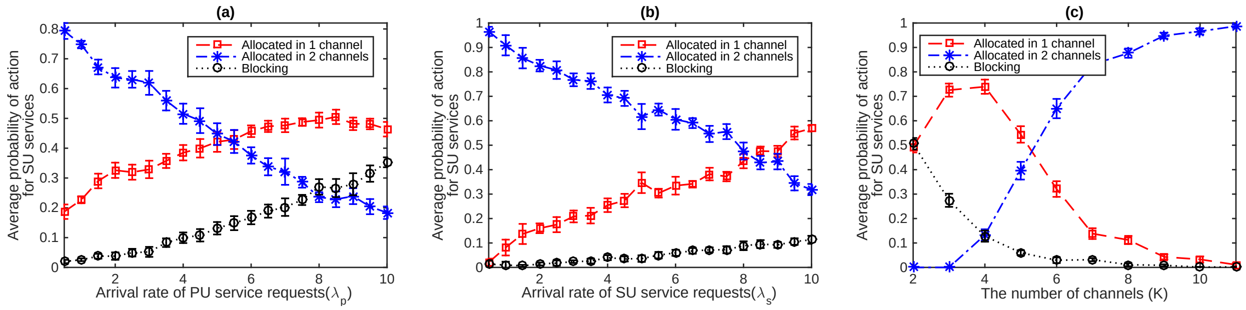


Fig. 3. The action probabilities of SU requests.

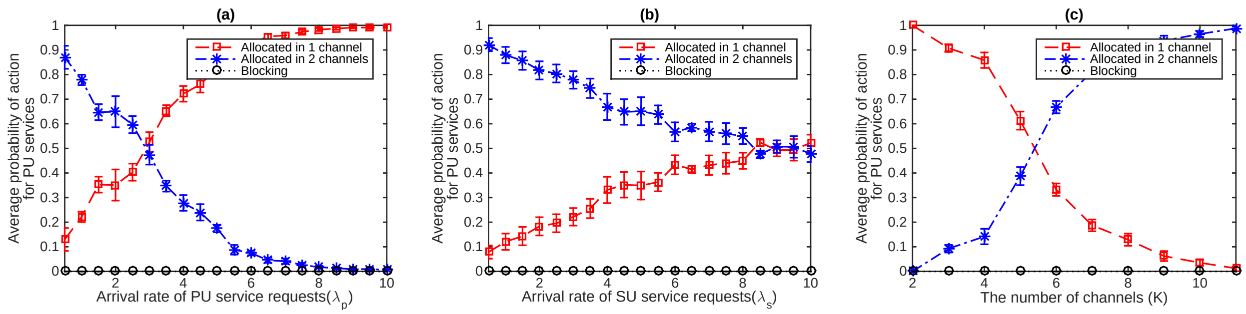


Fig. 4. The action probabilities of PU requests.

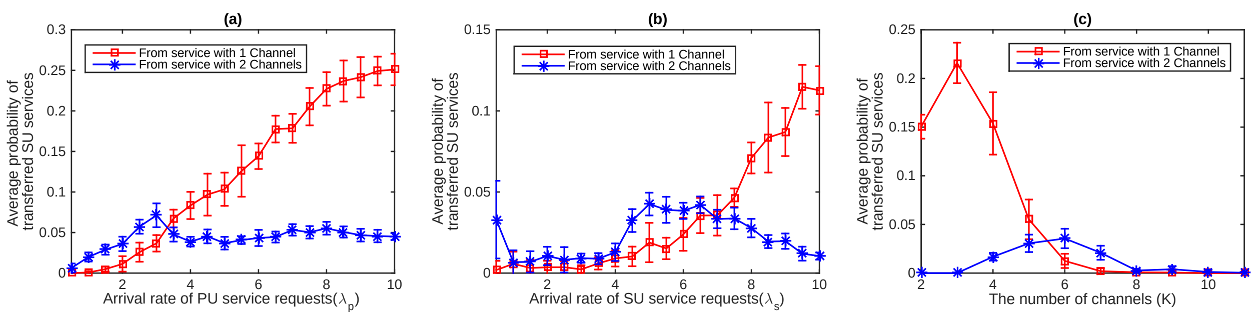


Fig. 5. The transfer action probabilities of SU requests.

long-term system reward of our model with the Greedy policy are shown in Fig. 6. We see the rewards of our model is much higher than the rewards by the Greedy policy, especially in the scenario of the channel resource shortage. When PU arrival rate λ_p increases in Fig. 6(a), the overall system reward raises continuously until the cost of transfer and blocking action influences the system. As the results of that, the rewards decrease when PU services arrival rate reaches to a certain value. When SU arrival rate λ_s is increasing in Fig. 6(b), the rewards of our policy are almost exponential growth since the PU arrival rate is fixed, while the rewards of the Greedy policy exhibit as a parabolic curve. The improvement in the overall system reward is also shown when the channel number K increases in Fig. 6(c). Furthermore, Fig. 7 shows the comparison in the PU service blocking rate. We see that the Greedy policy is not able to control the rejection rate when the arrival rate is high, or the number of channels is low, while our policy always maintains the rejection rate for PU to be zero.

VI. Conclusions

In this paper, we proposed a channel allocation scheme in a cognitive enabled VANET network based on SMDP policy. We analyzed the reward model of one roadside unit scenario with elements of users satisfaction and QoS requirement. The overall system long-term reward is maximized through adaptively allocating an appropriate number of channels to different categories of service requests. Also,

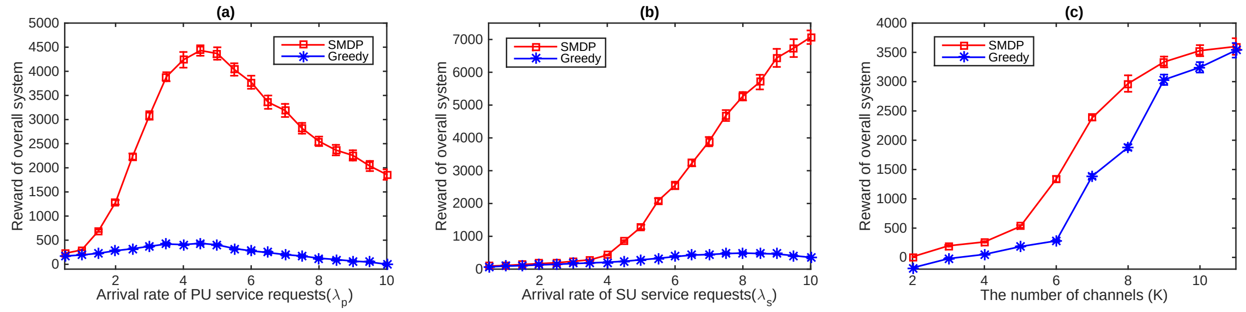


Fig. 6. The unified overall system rewards under different policies.

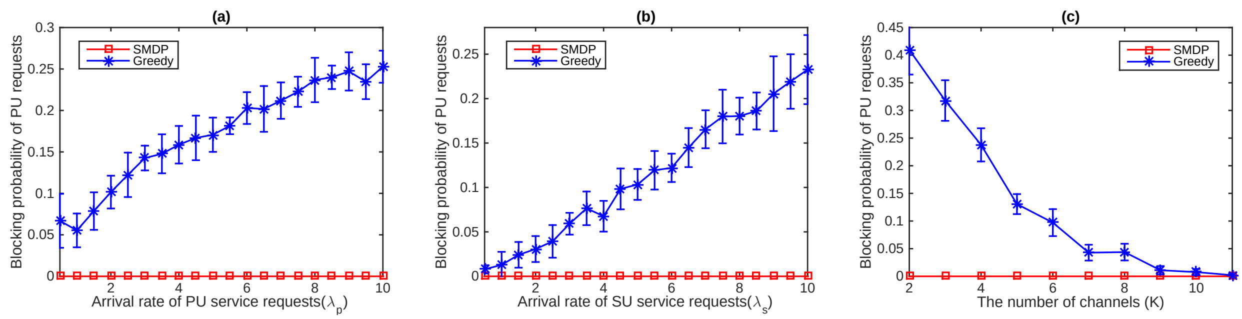


Fig. 7. The blocking probability of PU requests under different policies.

requests from primary users are given a higher priority to access the channel in our policy. For future works, the influence of interference between SU and PU channels in channel allocation process needs to be investigated in our model. Moreover, an alternative reward scheme needs to be studied considering diverse traffic environment, such as the direction and speed for vehicles.

Acknowledgement

The authors sincerely acknowledge the support from Natural Sciences and Engineering Research Council (NSERC) of Canada under Grant number RGPIN-2014-03777, the National Natural Science Foundation of China (NSFC) under Grant number 61571375 and the National High-Tech Research and Development Program of China (863 Program) under Grant number 2015AA01A705. Hongbin Liang is the corresponding author.

References

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

[2] T. H. Luan, L. X. Cai, J. Chen, X. S. Shen, and F. Bai, "Engineering a distributed infrastructure for large-scale cost-effective content dissemination over urban vehicular networks," *IEEE Trans. Vehicular Tech.*, vol. 63, no. 3, pp. 1419–1435, Mar 2014.

[3] G. Charalampopoulos, T. Dagiuklas, and T. Chrysikos, "V2I applications in highways: How RSU dimensioning can improve service delivery," in *2016 23rd International Conference on Telecommunications (ICT)*, May 2016, pp. 1–6.

[4] Y. Zhuang, J. Pan, Y. Luo, and L. Cai, "Time and location-critical emergency message dissemination for vehicular ad-hoc networks," *IEEE J. Select. Areas Communications*, vol. 29, no. 1, pp. 187–196, Jan 2011.

[5] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," *IEEE Signal Processing Magazine*, vol. 24, no. 3, pp. 79–89, May 2007.

[6] H. He, H. Shan, A. Huang, and L. Sun, "Resource allocation for video service in heterogeneous cognitive vehicular networks," *IEEE Trans. Vehicular Tech.*, vol. PP, no. 99, pp. 1–1, 2016.

[7] A. Alshamrani, X. S. Shen, and L. L. Xie, "QoS provisioning for heterogeneous services in cooperative cognitive radio networks," *IEEE J. Select. Areas Communications*, vol. 29, no. 4, pp. 819–830, Apr 2011.

[8] X. Y. Wang and P. H. Ho, "A novel sensing coordination framework for CR-VANETs," *IEEE Trans. Vehicular Tech.*, vol. 59, no. 4, pp. 1936–1948, May 2010.

[9] Y. Sun and K. R. Chowdhury, "Enabling emergency communication through a cognitive radio vehicular network," *IEEE Communications Magazine*, vol. 52, no. 10, pp. 68–75, Oct 2014.

[10] L. Zhang, T. Luo, W. Liu, S. Zhu, and J. Li, "Cooperative spectrum allocation with QoS support in cognitive cooperative vehicular ad hoc networks," *China Communications*, vol. 11, no. 10, pp. 49–59, Oct 2014.

[11] N. Zhang, H. Liang, N. Cheng, Y. Tang, J. W. Mark, and X. S. Shen, "Dynamic spectrum access in multi-channel cognitive radio networks," *IEEE J. Select. Areas Communications*, vol. 32, no. 11, pp. 2053–2064, Nov 2014.

[12] S. Kaplan, M. A. Guvensan, A. G. Yavuz, and Y. Karalurt, "Driver behavior analysis for safe driving: A survey," *IEEE Trans. Intellegent Transportation System*, vol. 16, no. 6, pp. 3017–3032, Dec 2015.

[13] A. A. Hammad, T. D. Todd, G. Karakostas, and D. Zhao, "Downlink traffic scheduling in green vehicular roadside infrastructure," *IEEE Trans. Vehicular Tech.*, vol. 62, no. 3, pp. 1289–1302, Mar 2013.

[14] T. Jiang, Z. Wang, L. Zhang, D. Qu, and Y. C. Liang, "Efficient spectrum utilization on TV band for cognitive radio based high speed vehicle network," *IEEE Trans. Wireless Communications*, vol. 13, no. 10, pp. 5319–5329, Oct 2014.

[15] N. Cheng, N. Zhang, N. Lu, X. Shen, J. W. Mark, and F. Liu, "Opportunistic spectrum access for CR-VANETs: A game-theoretic approach," *IEEE Trans. Vehicular Tech.*, vol. 63, no. 1, pp. 237–251, Jan 2014.

[16] X. L. Huang, J. Wu, W. Li, Z. Zhang, F. Zhu, and M. Wu, "Historical spectrum sensing data mining for cognitive radio enabled vehicular ad-hoc networks," *IEEE Trans. on Dependable and Secure Computing*, vol. 13, no. 1, pp. 59–70, Jan 2016.

[17] M. H. Cheung, F. Hou, V. W. S. Wong, and J. Huang, "DORA: Dynamic optimal random access for vehicle-to-roadside communications," *IEEE J. Select. Areas Communications*, vol. 30, no. 4, pp. 792–803, May 2012.

[18] L. Sun, H. Shan, A. Huang, L. Cai, and H. He, "Channel allocation for adaptive video streaming in vehicular networks," *IEEE Trans. Vehicular Tech.*, vol. PP, no. 99, pp. 1–1, 2016.

[19] C. Jiang, Y. Chen, and K. J. R. Liu, "Data-driven optimal throughput analysis for route selection in cognitive vehicular networks," *IEEE J. Select. Areas Communications*, vol. 32, no. 11, pp. 2149–2162, Nov 2014.

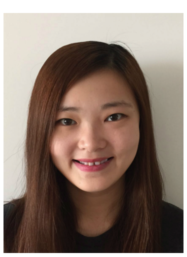
[20] K. Zheng, H. Meng, P. Chatzimisios, L. Lei, and X. Shen, "An SMDP-based resource allocation in vehicular cloud computing systems," *IEEE Trans. Ind. Electron.*, vol. 62, no. 12, pp. 7920–7928, Dec 2015.

[21] H. Liang, L. X. Cai, D. Huang, X. Shen, and D. Peng, "An SMDP-based service model for interdomain resource allocation in mobile cloud networks," *IEEE Trans. Vehicular Tech.*, vol. 61, no. 5, pp. 2222–2232, Jun 2012.

[22] X. Xiong, L. Hou, K. Zheng, W. Xiang, M. S. Hossain, and S. M. M. Rahman, "SMDP-based radio resource allocation scheme in software-defined internet of things networks," *IEEE Sys. J.*, vol. 16, no. 20, pp. 7304–7314, Oct 2016.

[23] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, Wiley, New York, 2005.

[24] R. L. Rivest C. Thomas, H. Cormen and C. Stein, *Introduction to algorithms*, MIT press, 2009.



Mushu Li obtained her Bachelor Degree from University of Ontario Institute of Technology (UOIT) in 2015. She is currently pursuing her MASci. degree at Ryerson University. She has been awarded OGS (Ontario Graduate Scholarship) in 2015 and 2016 respectively, and Aditya Jha Diversity Award from Ryerson University to recognize her academic achievement in 2016. Her research interests are the system optimization in VANETs, load balancing for smart grid, and EV charging scheduling.

Lian Zhao (S99-M03-SM06) received the Ph.D. degree from the Department of Electrical and Computer Engineering (ELCE), University of Waterloo, Canada, in 2002. She joined the Department of Electrical and Computer Engineering at Ryerson University, Toronto, Canada, in 2003 and a Professor in 2014. Her research interests are in the areas of wireless communications, radio resource management, power control, cognitive radio and cooperative communications, optimization for complicated systems.

She received the Best Land Transportation Paper Award from IEEE Vehicular Technology Society in 2016; Top 15 Editor in 2015 for IEEE Transaction on Vehicular Technology; Best Paper Award from the 2013 International Conference on Wireless Communications and Signal Processing (WCSP) and Best Student Paper Award (with her student) from Chinacom in 2011; the Ryerson Faculty Merit Award in 2005 and 2007; the Canada Foundation for Innovation (CFI) New Opportunity Research Award in 2005, and Early Tenure and promotion to Associate Professor in 2006. She has been an Editor for IEEE TRANSACTION ON VEHICULAR TECHNOLOGY since 2013; workshop co-chair for IEEE/CIC ICCC 2015; local arrangement co-chair for IEEE VTC Fall 2017 and IEEE Infocom 2014; co-chair for IEEE Global Communications Conference (GLOBECOM) 2013 Communication Theory Symposium.

She served as a committee member for NSERC (Natural Science and Engineering Research Council of Canada) Discovery Grants Evaluation Group for Electrical and Computer Engineering since 2015; an Associate Chair at the Department of Electrical and Computer Engineering at Ryerson University 2013-2015. She is a licensed Professional Engineer in the Province of Ontario, a senior member of the IEEE Communication and Vehicular Society.

Hongbin Liang (M12) received the B.Sc. degree in communication engineering from Beijing University of Post and Telecommunication, Beijing, China, in 1995, and the M.Sc. and Ph.D. degree in electrical engineering from Southwest Jiaotong University, Chengdu, China, in 2001 and 2012 respectively. He is currently an associate professor in the School of Transportation and Logistics, Southwest Jiaotong University. From 2001 to 2009, he was a Software Engineer with the Motorola R&D Center of China, where he focused on system requirement analysis and Third-Generation Partnership Project protocol analysis. From 2009 to 2011, he was a visiting Ph.D. student with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada. His current research interests focus on resource allocation, quality-of-service, security and efficiency in cloud computing, and wireless sensor networks.