Learning-Based Channel Selection of VDSA Networks in Shared TV Whitespace

Si Chen*, Rama Vuyyuru†, Onur Altintas‡, and Alexander M. Wyglinski*

*Wireless Innovation Laboratory, Electrical and Computer Engineering, Worcester Polytechnic Institute
Worcester, MA 01609, USA, Email: {chensi, alexw}@ece.wpi.edu

†Toyota InfoTechnology Center USA, Mountain View, CA 94043, USA, Email: rama@us.toyota-itc.com

‡Research & Development Division, Wireless Communications Group, Toyota InfoTechnology Center, Co., Ltd.
Tokyo, 107-0052, Japan, Email: onur@jp.toyota-itc.com

***Abstract*—In this paper, we propose a reinforcement learning-based approach for enabling vehicles to make intelligent channel selection choices across TV whitespace spectrum. In order for vehicle communication networks to dynamically access TV whitespace in a secondary manner, it is imperative that these communication systems be capable of coexisting with other types of secondary wireless networks operating within the same frequency range. Consequently, we first propose a TV whitespace channel sharing scheme that would facilitate the coexistence between WLAN, WRAN, and vehicular communication networks. Using the channel utilization variations observed by a collection of mobile vehicular communication systems, we then devised a reinforcement learning-based adaptive channel selection algorithm that employs channel utilization sensing in order to reinforce the decisions made by the vehicular communication system. Moreover, the parameters of the proposed learning approach are adaptively tuned in order to achieve better adaptation to a particular environment. A computer emulation environment composed of actual real-world sensing measurement data and a simulated TV whitespace network is created in order to accurately model the characteristics of future wireless environment, as well as to test the proposed learning-based channel access approach. Experimental results show a significant performance improvement with respect to vehicle communication.**

I. Introduction

In order to enhance the driving experience, especially with respect to increasing driver awareness and situational perception in order to ensure overall vehicular safety, a growing number of vehicles are equipped with wireless transmission systems to support either vehicle-to-infrastructure (V2I) or vehicle-to-vehicle (V2V) information exchanges. Several forms of V2I communications include but are not limited to vehicular communication systems accessing mobile cellular networks. For instance, as the number of smart handsets multiplies, telecommunications companies have deployed services that can provide Internet access supporting mobility up to 350 km/h in major cities. V2V communications are another important type of vehicle communications that provide ad hoc communications among vehicles within a region. V2V communications can facilitate greater vehicle coordination in order to prevent collisions and improved vehicle "platooning" capabilities. Moreover, the requirements on short delay and frequent ad hoc connections that exist with these forms of applications exclude cellular networks as a feasible candidate. Consequently, the Federal Communications Commission (FCC) allocated a 75 MHz band in the 5.9 GHz band for dedicated short-range communications (DSRC) [1], which is specially designed for vehicle communications.

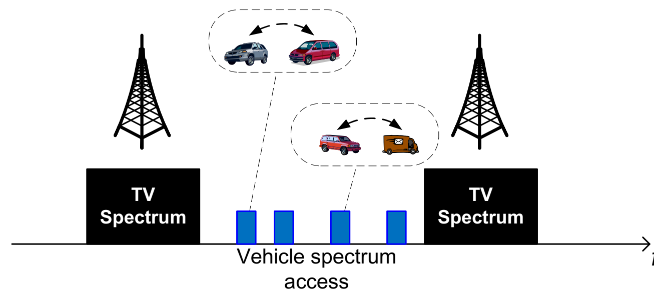

Fig. 1. Vehicular dynamic spectrum access in TV whitespace

However, given the finite amount of RF spectrum and the rapidly growing number of wireless applications and end-users, many sectors of wireless communication including the automotive sector are experiencing the effects of *spectrum scarcity*. DSRC channels are likely to face the same scarcity when vehicle communications becomes more popular. Conversely, it has been shown that the radio frequency spectrum is vastly underutilized [2]. To help alleviate this spectrum scarcity issue in order to accommodate the constantly growing demand for RF spectrum and greater wireless access, *dynamic spectrum access* (DSA) is considered by many to be the primary solution to this increasingly important problem. In order to realize DSA transmissions, *cognitive radio* can be used since it employs autonomous and flexible communications techniques implemented via a *software-defined radio* (SDR) programmable wireless communications platform. Furthermore, cognitive radio systems are often designed to be environmentally aware, which is essential when operating in transmission conditions involving licensed signals that can emerge at any time.

Although several licensed frequency bands could facilitate DSA-based vehicular communication networks, the ultra-high frequency (UHF) television frequency range which spans across 470-698 MHz (Channel 14 through Channel 51) has often been identified as a primary candidate due to its relatively

978-1-4673-1881-5/12/$31.00 ©2012 IEEE

static but spatially heterogeneous frequency channel utilization by incumbent TV broadcasters. In fact, the recent decision by the FCC to permit the wireless access of licensed DTV spectrum by secondary users [3] has made the concept of *vehicular dynamic spectrum access* (VDSA) a viable option for vehicular communication networks (see Fig. 1). However, there exists several other standards that are also being developed for operation in TV whitespace (TVWS), including ECMA 392 Standard, IEEE 802.11af Draft Standard, and IEEE 1900.x working groups, as well as one standard that has been recently ratified, namely, the IEEE 802.22 Standard [4]. With the decision by the FCC to allow secondary access networks in TV spectrum, several new technical challenges have emerged, including the issue of supporting multiple non-cooperative secondary wireless access networks across the same frequency range. Specifically, it is unclear how a VDSA network can potentially coexist with other secondary wireless access systems that are compliant with standards such as IEEE 802.22 and IEEE 802.11af.

In this paper, we propose a reinforcement learning-based approach for enabling vehicles to make intelligent channel selection choices across TV whitespace spectrum given the presence of other unrelated secondary wireless access networks. First, a TV whitespace channel sharing scheme is devised that would facilitate the coexistence between different secondary wireless access networks, including vehicular communication networks based on VDSA. Then, using the channel utilization variations observed by a collection of mobile vehicular communication systems, we then implement a reinforcement learning-based adaptive channel selection algorithm that employs channel utilization sensing such that decisions made by the vehicular communication system are reinforced with this information. Finally, the parameters of the proposed learning approach are adaptively tuned such that the results are tailored to a particular operating environment.

The rest of this paper is organized as follows: Section II describes the channel access problem in vehicle communications. Section III describe a proposed reinforcement learning based channel selection algorithm capable of adapting to the spectrum sharing environment. The algorithm is evaluated in a semi-realistic simulation environment in section IV. Section V concludes the paper.

II. Channel Access in Dynamic Vehicular Environments

In this section, we describe the available channels for vehicular communications and their access schemes. This is tightly connected to the issue of coexistence and sharing in channels where vehicular communications do not have exclusive authority. At this moment, there are three types of channels that are currently or potentially available to vehicle communications, including the dedicated channels such as the 5.9 GHz DSRC channels in North America and 700 MHz ITS band in Japan, the ISM bands that are open to free access, and opportunistic access channels such as TV whitespace. We are primarily concerned with V2V communications, and the implementation of V2I communications and mobile cellular networks are excluded from this discussion. Throughout, we refer all vehicle communication sets that need to access the wireless medium as vehicle groups, each containing one or many vehicles.

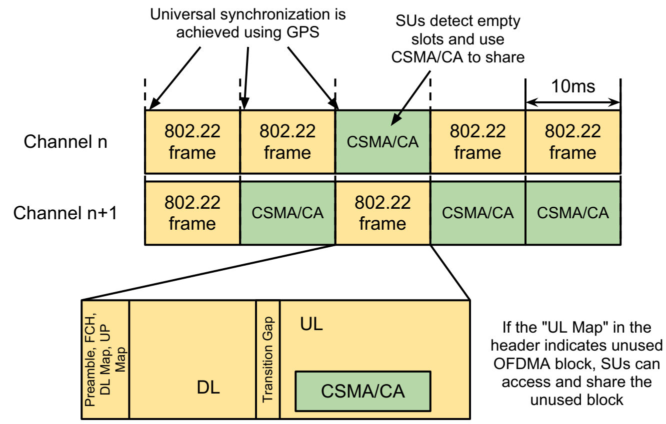

Fig. 2. An example of coexistence between frame-based PHY and contention-based PHY

Opportunistic access channel model has only existed in theory until recent when TV whitespace was opened to secondary wireless access. This could be a viable option for vehicle communication and an example of future flexible spectrum regulation. For instance, we have studied the feasibility of using TV whitespace for vehicle communications in [5]. However, the primary form of networks that are going to take advantage of TV whitespace at present is not vehicular networks but WRANs and WLANs, which are composed of fixed devices and portable devices of low mobility. Similar coexistence scenario is similar to the once expected scenario in 5.8 GHz U-NII/ISM license-exempt band, where 802.16 networks could share the same spectrum with 802.11a networks. It was suggested in [6] that coexistence causes a greater problem on 802.16 networks than on 802.11a networks, since 802.16 networks need rigorous protection to maintain the frame integrity, especially the frame header. In addition, IEEE 802.19 Task Group 1 was established targeting at the coexistence of networks using different standards [7].

In order to achieve coexistence between vehicle groups and WRANs, V2V communications can benefit from either one-way or two-way coordination with WRAN base stations, such as IEEE 802.22 networks. The synchronization required by such coordination can be achieved at vehicles by synchronizing to a global time reference, such as GPS. Since each 802.22 base station would be armed with a GPS device to report its location and perform channel sharing with other 802.22 base stations, vehicle networks and 802.22 networks already possess the capability of synchronous coexistence. Fig. 2 shows an example of frame-based networks and contention-based networks sharing TVWS channels.

A. Vehicular Dynamic Spectrum Access

Deploying a VDSA network is especially challenging due to the potential for a rapidly mobile communications environment, e.g., highway conditions. As a result, DTV spectrum has been identified as one candidate frequency band that can potentially support VDSA since the spectrum occupancy characteristics of the primary transmissions, i.e., DTV broadcast signals, are relatively static in both frequency and time. On the other hand, given the mobility aspects of a vehicular communication network in general, the spectrum occupancy characteristics across a range of frequencies starts becoming a time-varying phenomenon due to vehicles either moving into or out of range of a DTV broadcasting site. Nevertheless, this change in the spectral occupancy characteristics as a function of mobility is gradual enough to allow for the vehicular wireless devices to spectrally sense the frequency locations of the DTV broadcasts, as well as verify the sensing measurements against a spectrum occupancy database, and take the appropriate actions, such as dynamically change the transmission band in order to avoid interference with the TV signals. Furthermore, given the predictable, well-defined transmission characteristics of DTV broadcasts, it is possible to characterize the behavior of this frequency band, thus allowing for VDSA communications that do not interfere with the incumbent TV broadcasts.

Although several types of channels can be employed in vehicle communications, they exhibit different and stochastic channel characteristics and availability, which poses the problem of channel selection for vehicle communications. The average car lifespan at present is 12 years, within which the characteristics of wireless environment will definitely change, including the spectrum regulations and channel utilization statistics. Since the wireless communications among vehicles are expected to be fully automated with minimum driver distraction, it is necessary for the vehicle radio devices to adapt to the environment change automatically.

III. Channel Selection

A learning architecture was proposed in [8] to achieve intelligence in vehicle communications, and the channel selection problem can be solved using reinforcement learning embedded in that architecture. In this section, we consider the coexistence scenario comprising four typical kinds of networks, which are DTV broadcasting networks as primary users, IEEE 802.22 networks as typical WRANs, IEEE 802.11af networks as typical WLANs, and vehicle dynamic channel accessing networks. As this moment, we assume the vehicle networks are using a contention-based access protocol.

A. Channel Capacity

While the frequency of a channel affects the path loss, hence determines the feasible applications, the utilization of a channel determines the instant availability and possible reward for a vehicular network. The channel utilization depends on the coexistence scheme of all wireless users. Here we assumed a general coexistence scheme illustrated in Fig. 2.

We can model the possible reward in each channel as a function of its channel properties. Then the reward value can serve as the reinforcement signals for the reinforcement learning algorithm. The properties of a shared channel include the bandwidth, noise level, percentage of 802.22 frames (P_{Frame}), average 802.22 frame utilization (P_{Util}), percentage of channel busy time in a CSMA/CA (ρ), which has a upper bound of ρ_{MAX}, and number of contention users in CSMA/CA duration (N_C).

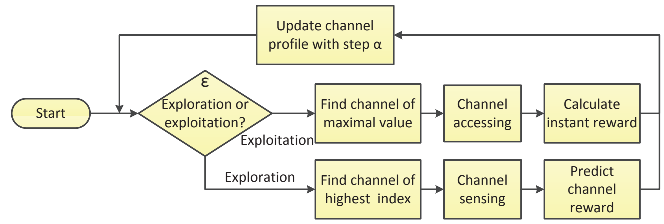

Fig. 3. The adaptive learning-based channel selection process.

The total capacity of a channel C can be determined by the bandwidth and noise level. We will focus on the remaining channel capacity C^R, which we model as:

$$C^R = C \cdot (1 - P_{\text{Frame}} \cdot P_{\text{Util}}) \cdot \left(\frac{\rho}{N_C + 1} + (\rho_{\text{MAX}} - \rho) \right) \quad (1)$$

The second term represents the temporal space left by the 802.22 WRANs. The third term is an indicator of the reward a new user can obtain via a fair contention scheme.

B. Channel Sensing

While it is straightforward to select a channel when the exact available channel capacities are known, the temporal variance in channel utilization and channel condition prevent a vehicle network to have perfect measurements of channel properties. Measuring channel utilization is more difficult compared with detecting signal existence, since the channel utilization also depends the accessing protocols used by coexisting users. At this point, we assume that a vehicle is only capable of sensing one channel at a time.

By synchronizing to the 10-ms 802.22 frames, a sensing period start with sensing 802.22 frame preamble and frame header. If no frame during that period, the sniffing mode for 802.11 packets will start, followed by channel idle time calculation. Otherwise, we decode the frame control header for uplink map to find OFDMA holes. If there exist OFDMA holes allowing contention-based secondary access, packet sniffing during OFDMA holes will start.

We found via analyzing the sniffing data of real WLANs that the channel utilization of a pure CSMA/CA channel follows a loglogistic distribution at low contention level and a normal distribution at high contention level. Similar statistics can be expected in future WLANs in TVWS as well.

C. Adaptive Learning Approach

In order to enable learning from past experience, a vehicle needs to keep a set of channel profiles for all possible channels. The profile of a channel s consists of a channel value $V(s)$, an index $I(s)$, and $n(s)$, the times of using channel s. Given $R(s)$, the sampled remaining channel capacity C^R on channel

s, the channel value $V(s)$ is updated using a Monte Carlo method [9]:

$$V(s) \leftarrow V(s) + \alpha(R(s) - V(s)), \quad (2)$$

where α is the learning rate between 0 and 1.

The channel learning process is shown in Fig. 3. The exploitation process consists of a sensing period followed by an accessing period, while the exploration process only does channel sensing. This arrangement is to avoid high time consumption caused by frequent channel switching. During the exploitation process, the channel with the maximal channel value V is selected as the access channel. During the exploration process, a vehicle will select a channel other than the exploitation channel to perform channel sensing for T_{Sensing}.

The objective of the exploration process here is to quickly find the best channel other than the current exploitation channel, such that the transmission can quickly switch to the new channel when the current access channel becomes unavailable. Hence the reward signal of channel sensing represents the predicted available capacity remained on that channel and the exploration scheduling algorithm will maximize the total reward. We use an index-based scheme to choose an exploration channel s with maximal channel index $I(s)$ defined as:

$$I(s) = V(s) + \sqrt{\frac{2\ln\left(\sum_{1<j<K} n(j)\right)}{n(s)}}, \quad (3)$$

where K is the total number of available channels. This is similar to the upper confidence bound (UCB) algorithm proposed in [10], where $V(s)$ is replaced by the average of reward received on channel s. This approach starts with sweeping the channel once and bias the channel selection toward those that have been sensed free in the near past. While a random channel selection assigns equal probabilities to all channels, which will waste the limited sensing time on highly utilized channels, our channel exploration scheduling approach will spend more sensing time on channels with low utilization.

In order to better understand and adapt to the environment, key parameters of the learning algorithm are adaptively tuned. The two key parameters are the exploration rate ε and the learning rate α, also known as the step-size. In our channel selection problem, since the learning agent on a vehicle is more interested in whether to switch to a new accessing channel or to remain in the current channel s, the exploration rate ε should represents the confidence level that the there is another channel that has higher remaining capacity than the current channel reward. Hence, ε can be defined as the probability that there exists another channel with greater remaining capacity than the current channel reward times the maximum possible improvement on channel reward. Assuming the channels are independent of each other, we have:

$$\varepsilon = \left(1 - \prod_{j=1, j\neq s} \Pr\left[{C^R}_j < {C^R}_s\right]\right) \cdot \left(1 - \frac{C^R}{C}\right). \quad (4)$$

When the current channel reward is high, the exploration rate ε is reduced. When exploration rate is reduced, the learning agent becomes insensitive to environment changes. In order to maintain sensitivity to channel variation, we also let the update step size α change inversely proportional to exploration rate ε.

IV. Evaluation of VDSA

In this section, we show the evaluation of VDSA in future TV whitespace. Our goal is to study the process of spectrum selection in a flexible spectrum access environment where multiple heterogeneous wireless networks share the same spectrum band. At this point, we do not limit this work to any specific applications or wireless standards, hence we are not looking into how users within a network interact with each other but evaluating the process of channel selection that all applications are build upon.

A. Simulation Setup

In the following simulations, we use the real channel measurements of UHF channel 21 to 51 across I-90 in Massachusetts taken in 2009 [11], and overlap artificial WRANs and WLANs onto the measurements. The following parameters are used for generating WRANs and WLANs:

	WRANs	WLANs
Density	1/6km	10/6km
Range	25km	2km
Mean of utilization	0.4	$0.1 \sim 0.5$
Variance of utilization	0.12	0.12

Each vehicle group is assumed to have a moving speed randomly generated between 20 m/s and 40 m/s. We tested a range of vehicle densities with the average number of vehicle groups in each direction per kilometer varying from 1 to 5. 30 scenarios were generated with WRANs and WLANs with random locations and frequency channel. WRANs are assumed to each occupy a DTV channel of 6 MHz, and each WLAN occupies two contiguous DTV channels. Each scenario is tested with vehicle densities from 1 to 5 for 1000 times with random vehicle flows.

B. Simulation Results

The following three approaches are compared with each other:

Random greedy approach: While continuously accessing the channel, the agent will take a moving average of the channel reward in the past one second, which is composed of ten samples of 100 ms slots, and see if the averaged normalized reward is less than 50% of the channel capacity. When the average is below 50%, the agent will jump to the first channel with less than 50% channel utilization via random sensing over all available channels.

Classic value iteration reinforcement learning: A classic value iteration approach is employed with adaptive exploration rate ε and learning rate α. The exploration channel is chosen randomly among all available channels.

Value iteration with UCB controlled exploration: Instead of a random channel exploration, the agent uses a UCB algorithm described in Section III to schedule channel exploration. Same

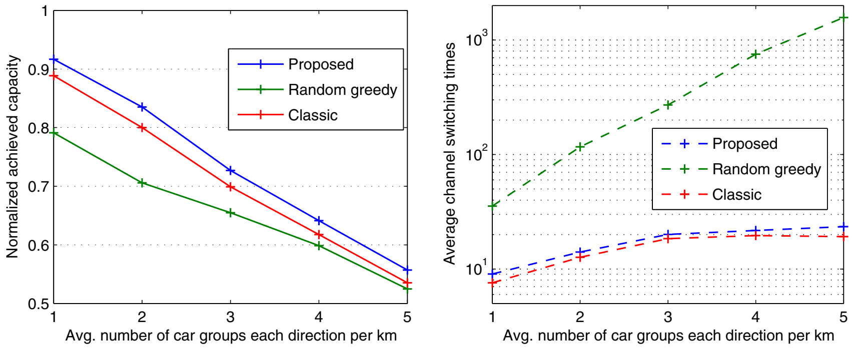

Fig. 4. (a) Normalized throughput, (b) Average switching times

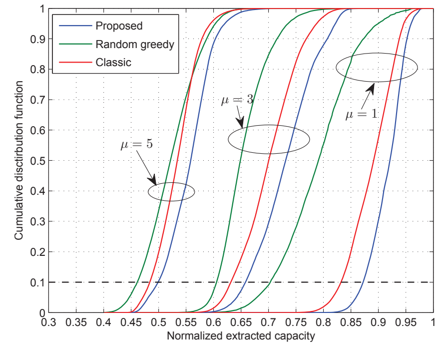

Fig. 5. CDF of achieved throughput. μ is the average number of car groups in each direction/km.

adaptation techniques in the exploration rate and step size are used.

Three metrics are used to compare the performance of the above approaches. The normalized throughput is calculated as the total achieved channel reward divided by the maximum channel capacity of an empty channel. The channel switching time cost, which can be as long as 2 seconds [12] depending on other implementation issues, is not deducted from the calculation of normalized throughput. Nevertheless, combining the results of normalized throughput and average channel switching times, the reader will be able to get a better understanding of the actual achievable throughput. The average number of channel switching times counts the times that an agent switch its accessing channel but not its sensing channel (exploration channel). Unlike switching the communication channel, the switching of a sensing channel can be accomplished instantly. The major time consumption in a network is due to switching the communication channel. The higher the channel switching times, the lower the time ratio for actual communications. The last metric is outage probability, which measures the probability of the achieved total channel reward falling below a certain threshold. This metric reflects the distribution of the achieved throughput over multiple runs.

Fig. 4 shows the comparison results of above three metrics. While the throughput of all three approaches decline as the vehicle traffic increases, adaptive approaches always perform better and UCB controlled channel exploration performs the best. The difference between the naive approach and the adaptive approaches is significant in terms of channel switching times. While the channel switching times of the naive greedy approach increases exponentially as vehicle density increases, the adaptive approaches can maintain an acceptable switching times. The slightly higher number of channel switching times of the UCB controlled channel exploration may be caused by the better capability of finding a better channel compared to the random channel exploration.

The cumulative distributions of the achieved throughput of all three approaches are given in Fig. 5. Three different traffic levels are employed with μ, the average number of car groups in each direction per kilometer, taking on values of 1, 3, and 5. Comparing the CDFs with different traffic levels, we can see more significant advantages of applying learning to channel selection in more crowded traffic.

V. Conclusion

Vehicular dynamic spectrum access (VDSA) is one solution to the spectrum scarcity issue faced by vehicular communication networks. In this paper, a reinforcement learning-based adaptive channel selection algorithm is proposed that is designed to enhance the spectrum access performance of vehicles by improving throughput and reducing channel switching times. At the same time, this proposed approach also enables the coexistence of VDSA networks with other secondary wireless access networks operating within the same frequency range.

References

[1] Federal Communications Commission, "FCC 99-305," FCC Report and Order, October 1999.

[2] S. Pagadarai and A. Wyglinski, "A quantitative assessment of wireless spectrum measurements for dynamic spectrum access," in *4th International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM)*, Jun. 2009.

[3] Federal Communications Commission, "Code of Federal Regulations, Title 47, Part 15-Radio frequency devices, Subpart H Television Band Devices."

[4] IEEE WG802.22, "IEEE 802.22-2011 Standard for Cognitive Wireless Regional Area Networks (RAN) for Operation in TV Bands," 2011.

[5] S. Chen, A. M. Wyglinski, R. Vuyyuru, and O. Altintas, "Feasibility analysis of vehicular dynamic spectrum access via queueing theory model," in *IEEE Vehicular Networking Conference (VNC)*, Dec. 2010.

[6] L. Berlemann, C. Hoymann, G. Hiertz, and B. Walke, "Unlicensed operation of IEEE 802.16: Coexistence with 802.11 (a) in shared frequency bands," in *IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications*, 2006, pp. 1–5.

[7] T. Baykas, J. Wang, and S. Filin, "Standardization activities of IEEE 802.19 Task Group 1: wireless coexistence in the TV White Space," vol. 110, no. 398, pp. 25–30, 2011.

[8] S. Chen, R. Vuyyuru, A. M. Wyglinski, and O. Altintas, "On Optimizing Vehicular Dynamic Spectrum Access Networks: Automation and Learning in Mobile Wireless Environments," in *IEEE Vehicular Networking Conference (VNC)*, 2011.

[9] R. Sutton and A. Barto, *Reinforcement learning: An introduction*. Cambridge Univ Press, 1998, vol. 116.

[10] P. Auer, N. Cesa-Bianchi, and P. Fischer, "Finite-time analysis of the multiarmed bandit problem," *Machine learning*, vol. 47, no. 2, pp. 235–256, 2002.

[11] S. Pagadarai, A. Wyglinski, and R. Vuyyuru, "Characterization of vacant UHF TV channels for vehicular dynamic spectrum access," in *IEEE Vehicular Networking Conference (VNC)*, 2009, pp. 1–8.

[12] O. Altintas, M. Nishibori, T. Oshida, Y. Ihara, and M. Saito, "Demonstration of Vehicle to Vehicle Communications over TV White Space (Demo paper)," in *4th International Symposium on Wireless Vehicular Communications (WIVEC)*, 2011.