



Deep Reinforcement Learning to Improve Vehicle-to-Vulnerable Road User Communications in C-V2X

Andy Triwinarko^{1,2(✉)}, Zoubeir Mlika¹, Soumaya Cherkaoui³,
and Iyad Dayoub⁴

¹ Université de Sherbrooke, Sherbrooke, QC J1K2R1, Canada

² Politeknik Negeri Batam, Batam 29641, Indonesia
andy@polibatam.ac.id

³ Polytechnique Montréal, Montréal, QC, Canada
soumaya.cherkaoui@polymtl.ca

⁴ University Polytechnique Hauts-de-France, 59313 Valenciennes, France
iyad.dayoub@uphf.fr

Abstract. In this paper, we study the problem of optimizing the performance of vehicle-to-everything (V2X) using deep reinforcement learning techniques while sharing the spectrum between vehicle-to-infrastructure (V2I) links and vehicle-to-vulnerable road users (V2VRU) links in Cellular V2X (C-V2X). The objective is to protect VRU by improving the performance of V2VRU communications while maximizing the performance of V2I communications. Specifically, we formulate a spectrum sharing optimization problem with a two-objective function where the first objective is to improve the packet reception ratio (PRR) of VRU, whereas the second objective is to maximize the data rate of V2I communication links. To solve this challenging problem, we propose a deep reinforcement learning algorithm. A single agent controlling the vehicular network observes the environment and takes decisions accordingly by appropriately selecting the spectrum sub-bands and the transmission power levels. The simulation results show that the proposed scheme attains high performance compared to baseline solutions and solves the trade-off between maximizing the data rates of the vehicle users (V2I links) and improving the PRR of the V2VRU links.

Keywords: Vehicular communications · vulnerable road users · Deep Reinforcement · spectrum sharing · optimization

1 Introduction

Vehicular-to-everything (V2X) communication is considered one of the key pillars of future generations of wireless networks. It offers diverse services ranging from infotainment services to safety services, such as road safety, ubiquitous Internet access, traffic efficiency, etc. The 3rd Generation Partnership Project (3GPP) organization and other industries have already started working to provide V2X communication and improve its services [1–4].

In this paper, we design a spectrum sharing strategy for vehicular networks that includes vulnerable road users (VRU) and non-VRU that supports different connectivity such as vehicle-to-VRU (V2VRU) and vehicle-to-infrastructure (V2I) connectivity [5]. The V2VRU communication concerns the links among vehicles and VRU, whereas the V2I communication concerns the links between vehicles and infrastructures such as roadside units (RSU). As discussed in 3GPP for V2V [2], the V2VRU communication is established on a side-link radio interface (called PC5), but the V2I communication is established on a cellular radio interface (called Uu).

The use case proposed in this paper consists of supporting two different services: (1) application or infotainment services for high data rate V2X communication and (2) advanced driving for high-reliability V2X communication. Our supported use case aims to serve two types of vehicular users: non-VRU (or simply vehicle users) and VRU, such as pedestrians, cyclists, etc. VRUs are particularly important to protect since pedestrians account for over 21% of road fatalities and motorcycles, bicycles and scooters over 26% [5,6]. VRUs are generally unpredictable in their movements and are not always visible in a line of sight manner to other non-VRU, which makes the propagation characteristics of V2VRU communication different from classical V2V or V2I communication. To support this use case, we aim (1) to maximize the data rate of non-VRU (V2I communications) and at the same time (2) to improve the packet reception ratio (PRR) of VRU while allocating the shared spectrum between both types of road users. This paper chooses pedestrians as our VRU, so we use vehicle-to-pedestrians (V2P) as our V2VRU communication.

1.1 Related Works

Although many research efforts have been made to solve the resource allocation problem in V2V communication [7–12], only few papers discussed V2VRU communication to protect VRU, even though the number of road fatalities for VRU is high. The authors in [13] investigate the performance radio access technology (RAT) standards IEEE 802.11p, 802.11bd, 4G LTE-V2X and 5G NR-V2X for V2VRU communication. Each RAT complies with the safety applications requirements, with 20–100 ms latency and 0.5–700 Mbps throughput. We want to improve the performance of the V2VRU communication links; hence it can also improve the protection of VRU.

The authors in [14] proposed resource sharing as a multi-agent reinforcement learning problem, where multiple V2V links reuse the frequency spectrum occupied with the V2I links. Several vehicles connected with V2V links act as an agent that observes the environment and learns to improve the spectrum and power allocation. We were inspired by this approach for our V2VRU communications, where we need to consider that the VRU channel model is different from the V2V and V2I channel models.

In [15] the authors studied the problem of network slicing to optimize the resource allocation problem in V2X communication while optimizing the coverage area of vehicles, transmission power, and RB allocation. The objective of the

problem is to maximize the number of two types of packets: infotainment packets and safety packets. The authors proposed a multi-agent deep reinforcement learning approach to solve the problem. The difference between this work and ours is that we consider spectrum sharing as well as V2VRU communication.

The authors in [16] proposed the channel model for vehicle-to-pedestrians (V2P) communication in three main scenarios, i.e., 1) the vehicle to static pedestrian, 2) vehicle to moving pedestrian, and 3) vehicle to pedestrians with crowd shadowing scenarios. We used the findings of this paper to model the V2P channel in our work.

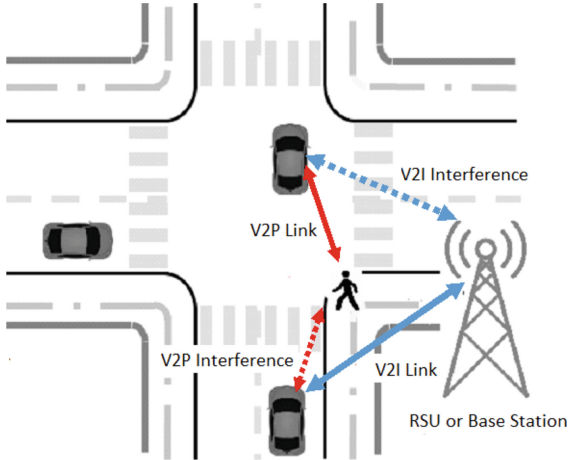


Fig. 1. An illustrative system of V2X networks consists of V2I and V2P links and its interference signal. The V2I links are preoccupied the spectrum bands and the V2P links will have a spectrum sharing with the V2I link.

2 System Model

We consider a vehicular network composed of a set of vehicles (non-VRU) and a pedestrian as our VRU. The vehicular network operates in a geographical zone covered by a cellular network (e.g., LTE or 5G), where a RSU is present to serve the road users. We consider two kinds of communications where vehicles communicate with the RSU (also known as V2I communication) and non-VRU that communicate with pedestrian (also known as V2VRU communication). We assume that there are m V2I links as well as n V2VRU links. The illustration of V2X communication using spectrum sharing between V2I and V2VRU (or V2P) links can be seen in Fig. 1.

The Mode 4 defined in cellular V2X communication is considered where all the vehicular users (VRU and non-VRU) share a pool of radio resources of V2I communication from which they must select appropriate resources to communicate between each other for the V2VRU communication. The overlap between

V2I resources and V2VRU resources can help in improving the spectral efficiency provided that the resources are optimally allocated. That is, interference caused by V2VRU should be carefully managed to improve the performance of both V2I and V2VRU communications.

As done in [14], we assume that the resources allocated to the V2I communication are assumed to be orthogonal among each other so that the interference is neglected in the V2I communication. In other words, V2I link $i = 1, \dots, m$ is assumed to communicate over the spectrum sub-band i . The main challenge is thus how to share the spectrum sub-band $i = 1, \dots, m$ between the V2I communication and the V2VRU so that the data rate of the V2I communication and the packet reception rate (PRR) of the V2VRU communication are maximized.

Note that each sub-band is defined simply as a set of consecutive sub-carriers that are grouped together. Here, orthogonal frequency division multiplexing (OFDM) is used as in [14] to transform the frequency selective channels into a set of parallel flat channels over multiple sub-carriers. It is further assumed that the channel fading is constant over the same sub-band but varies independently from one sub-band to another.

Once a V2VRU link j is allocated sub-band $i = 1, \dots, m$, its channel power gain is given by:

$$g_{j,j}^i = \text{PL}_{j,j} h_{j,j}^i, \quad (1)$$

where $\text{PL}_{j,j}$ is the large-scale fading that includes path-loss and shadowing propagation effects. It mainly depends on the distance of the V2VRU link j . The small-scale fading is denoted by $h_{j,j}^i$, which is frequency-dependent and generally assumed to follow an exponential distribution. However, in this work, the propagation characteristics of V2VRU link j is modeled differently to capture the effect of VRU.

The channel power gain over the V2I link i is denoted as $g_{i,0}^i$ (between the transmitter of link i and the RSU denoted by 0), which includes small-scale and large-scale fading as in (1). (The superscript i denotes the sub-band i allocated to V2P link i .)

There are three types of interference in our system model.

- V2I-to-V2VRU interference: The interference caused by V2I link i on the V2VRU link j (after allocating sub-band i to link j) is denoted by $I_{0,j}^i = p_i g_{0,j}^i$. Here, $g_{0,j}^i$ denotes the channel power gain between the transmitter of the V2I link i and the receiver of the V2VRU link j over sub-band i .
- V2VRU-to-V2VRU interference: The interference caused by other V2VRU link j' on V2VRU j when both links are allocated sub-band i is denoted by $I_{j',j}^i = p_{j'} g_{j',j}^i$. Here, $g_{j',j}^i$ denotes the channel power gain between the transmitter of the V2VRU link j' and the receiver of the V2VRU link j over sub-band i .
- V2VRU-to-V2I interference: The interference caused by V2VRU link j on the RSU when the V2VRU link j is allocated sub-band i is denoted by $I_{j,0}^i = p_j g_{j,0}^i$. Here, $g_{j,0}^i$ denotes the channel power gain between the transmitter of the V2VRU link j and the receiver of the V2I link i over sub-band i .

All channel power interference discussed previously are assumed to be defined similarly to (1). The model of our wireless channel of the V2I and the V2VRU are given respectively in [2] and [16].

On the one hand, the data rate of the V2I link i (over sub-band i) is given as:

$$R_i = B \lg(1 + \text{SINR}_i), \quad (2)$$

where B is the bandwidth of each sub-band and SINR_i is given in the following equation:

$$\text{SINR}_i = \frac{p_i g_{i,0}^i}{\sigma^2 + \sum_j x_j^i I_{j,0}^i}, \quad (3)$$

where the denominator in (3) contains the noise over sub-band i and the interference coming from all other V2VRU links j that are transmitting over sub-band i as well. The variables x_j^i are used to denote whether sub-band i is allocated to the V2VRU link j or not. The transmission power of the transmitter in the V2I link i is denoted as p_i (constant) and the transmission power of the transmitter in the V2VRU link j over sub-band i is denoted as p_j^i (variables).

On the other hand, the data rate of the V2VRU link j (over sub-band i , i.e., $x_j^i = 1$) is given as:

$$R_j^i = B \lg(1 + \text{SINR}_j^i), \quad (4)$$

where SINR_j^i is given in the following equation:

$$\text{SINR}_j^i = \frac{p_j^i g_{j,j}^i}{\sigma^2 + I_{0,j}^i + \sum_{j' \neq j} x_{j'}^i I_{j',j}^i}, \quad (5)$$

where the denominator in (5) contains the noise over sub-band i and the interference coming from all other V2VRU links j' that are transmitting over sub-band i as well as from the transmitter of the V2I link i . To protect VRU from accident and other fatalities, we have to guarantee a reliable communication for the safety-critical message exchanged over the V2VRU links. For this reason, we use the performance metric called packet reception rate (PRR) [14] which mainly measure how much safety packets are delivered during a defined time window. Mathematically, the PRR is defined as [14]:

$$\Pr\left[\sum_{t=1}^T \sum_{i=1}^m x_j^i R_j^i(t) \geq \gamma_j / \delta_T\right], \forall \text{ V2VRU link } j, \quad (6)$$

where T denotes a defined period of time during which the packet of the V2VRU link j should be transmitted and γ_j is the size of the safety packet of the V2VRU link j . The variable δ_T denotes the channel coherence time. We added the time index t to the notation of the data rate of the V2VRU link j over sub-band i , $R_j^i(t)$, to denote the data rate at each time instant t (each coherence time-slot). We used $\Pr[\cdot]$ to denote the probability function.

3 Problem Formulation

In this paper, the objective is to maximize the data rate of the V2I links as well as the packet reception rate of the V2VRU links. This problem is formulated as follows:

Problem 1.

$$\underset{\mathbf{x}, \mathbf{p}}{\text{maximize}} \quad \alpha \sum_i R_i + \beta \sum_{j=1}^n \Pr \left[\sum_{t=1}^T \sum_{i=1}^m x_j^i R_j^i(t) \geq \gamma_j / \delta_T \right] \quad (7)$$

$$\text{subject to} \quad x_j^i \in \{0, 1\}, p_j^i \geq 0 \forall i, j, \quad (8)$$

$$\sum_{i=1}^m x_j^i \leq 1, \forall j, \quad (9)$$

$$\sum_{i=1}^m p_j^i \leq \bar{p}_j, \forall j. \quad (10)$$

The variable \mathbf{x} is a matrix notation for the binary variables x_j^i for the sub-band allocation whereas the variable \mathbf{p} is a matrix notation for the real variables p_j^i for the power allocation. Constraints (9) guarantee that each V2VRU link is allocated at most one sub-band i . Constraints (10) guarantee that each V2VRU link j is allocated a maximum transmission power \bar{p}_j over all its allocated sub-bands. The parameters α and β are used to weigh the two-objective function to make it unit-less.

Problem (1) is challenging to solve due to the non-convexity of the objective function. To efficiently solve it, we propose a machine learning approach based on single-agent deep reinforcement learning (SARL).

4 Proposed Solution

In this section, we describe our proposed SARL approach. First, we describe the agent, the action and the state spaces, and the transition probability function. Then, we describe the algorithm.

The DRL agent is implemented in the RSU. This means that it is a centralized agent that observes the initially unknown vehicular environment and collects information to take actions accordingly. The agent exchanges information between non-VRU and VRU using dedicated wireless channels without excessive overheads.

The state space is roughly the vehicular network. More precisely, at each coherence time instant t , a state s_t is observed. The state s_t is generally unknown and includes channel conditions of all vehicles, vehicle movements, radio access information such as transmission power and spectrum allocation, etc. The DRL agent can extract useful information from this unknown state s_t through an observation function that maps each unknown state s_t into a well-defined and

known observation variable o_t . The observation o_t includes the channel power gains g_j^i , $g_{i,0}$ as well as the interference $I_{j,0}^i$, $I_{j',j}^i$, and $I_{0,j}^i$. It also includes the sum of interference illustrated in the SINR formulas in (3) and (5). Finally, we add to the observation o_t the variables γ_j^r and $T^r = T - t$ which denote the remaining number of bits of V2VRU's j packet and the remaining time interval. Note that, we can, for simplicity, exchange the s_t and the o_t notations.

The action space includes the sub-band spectrum allocation as well as the transmission power levels. We assume a set of P transmission power levels from which each V2VRU link can choose. Thus an action at time instant t is given by $\mathbf{a}_t = [(p_1(t), x_1(t)), \dots, (p_j(t), x_j(t)), \dots, (p_n(t), x_n(t))]^\top$ which is a vector of length n —the number of V2VRU links. Each element of the vector \mathbf{a}_t at time instant t is a pair $(p_j(t), x_j(t))$ of transmission power levels $p_j(t) \in \{1, 2, \dots, P\} \cup \{0\}$ and of sub-band allocation $x_j(t) \in \{1, 2, \dots, m\}$. The power level 0 indicates no transmission and $x_j(t) = i$ indicates that the V2VRU link j is allocated the sub-band i at time instant t .

The reward of the DRL agent is chosen to reflect the objective function in problem (1). That is, given the current state s_t , the reward of the DRL agent is given as $\eta_1 > 0$ or $0 < \eta_2 \ll \eta_1$. The value of η_1 (resp. η_2) denotes the number of V2VRU links that terminated transmitting (still transmitting) their safety packets. In this way, we encourage the DRL agent to accumulate as large reward as possible by serving V2VRU links and finishing successfully their safety packet transmissions.

The transition from one state to another is given by a probability distribution that gives how to transition from state s_t to state s_{t+1} , given action a_t , i.e. $\Pr[s_{t+1}, r_{t+1} | s_t, a_t]$ is the probability of moving to state s_{t+1} from state s_t and obtaining the reward r_{t+1} when taking action a_t . This probability distribution depends on the dynamical vehicular environment including the channel conditions and the vehicle mobility and it is generally hard to compute explicitly due to the complex nature of the vehicular environment.

The proposed algorithm is based on deep-Q-learning (DQL). DQL combines the well-known Q-learning method and deep neural networks. Q-learning creates a table of state-action pairs called the Q-table and finds the best action given a certain state using a greedy exploration approach, called ϵ -greedy. In ϵ -greedy, an action is chosen at random with probability ϵ and the action that gives the best reward is chosen otherwise. The drawback of Q-learning appears when the state and action spaces become large. Also, once the size of the table grows very large, many states will be very rarely visited, which deteriorates the learning strategy.

The key success of deep-Q-networks (DQN) is the use of experience replay memory technique where the tuple of state, action, reward, and next state are stored in a replay buffer. Next, the DRL agent samples from this replay buffer to perform learning. DQL is a promising approach that can be used to solve the curse of dimensionality in RL [15] by approximating the Q-table. We combine DQN with independent Q-learning.

Initially, the experience replay memory and all parameters are initialized. Then, the proposed algorithm SARL operates over a set of episodes. In each episode, the DRL agent explores the action space using the ϵ -greedy policy.

Table 1. Log-Distance Path Loss Parameters in V2I and V2P Based on Reference [14, 16]

V2X Types	Parameters	Values
V2I Link	Path loss	128.1 dB
	Path loss exponent	3.76
	Shadowing standard deviation	8 dB
	Break-point distance	50 m
V2P scenario 1 (static pedestrian)	Path loss	46.77 dB
	Path loss exponent	2.03
	Shadowing standard deviation	3.20 dB
V2P scenario 2 (moving pedestrian)	Path loss	40 dB
	Path loss exponent	2.44
	Shadowing standard deviation	5.47 dB
V2P scenario 3 (crowded pedestrian)	Path loss	67 dB
	Path loss exponent	1.26
	Shadowing standard deviation	3.35 dB

Each episode e covers a time horizon of T time-slots. At the beginning of the first time-slot, the starting state of the vehicular environment (initial positions of the vehicles, of the pedestrians, etc.) is revealed to the DRL agent. For subsequent time-slot t , the DRL agent chooses an action \mathbf{a}_t for each V2VRU link, according to the ϵ -greedy approach. In other words, the DRL agent chooses a transmission power level and a sub-band for each V2VRU link $j = 1, 2, \dots, n$. Once all V2VRU links have been allocated a transmission power level and a sub-band, the DRL agent evaluates the reward function r_{t+1} based on the expression of (2) and (4). Next, each V2VRU link (each vehicle and pedestrian) moves, according to its mobility model, and the next state is revealed to the DRL agent. The resulting tuple $(s_t, \mathbf{a}_t, r_{t+1}, s_{t+1})$ is collected and is stored in the prioritized experience replay memory of the DRL agent. This experience replay memory is associated some positive priority weight. After a few episodes, a mini-batch is sampled according to their priorities from the prioritized experience replay memory. This mini-batch is used to update the DQN weight parameters using a variant of the stochastic gradient descent algorithm to minimize the loss function. The loss function is given by the mean square error as follows:

$$\sum_{(s_t, \mathbf{a}_t, r_{t+1}, s_{t+1})} \left(r_{t+1} + \gamma \max_a (Q(s_{t+1}, a; \mathbf{w}^-)) - Q(s_t, \mathbf{a}_t; \mathbf{w}) \right)^2, \quad (11)$$

where the DQN is represented mathematically by the Q-function $Q(s_t, \mathbf{a}_t; \mathbf{w})$ (the function that the DQL tries to approximate) and \mathbf{w}^- is the weight parameter of a duplicate copy of the original DQN (called the target DQN) that is

created in order to update the original DQN from occasionally. The creation of a target DQN is suggested by the quasi-static target network method [17] to set the targets of the Q values.

5 Simulation Results

In this section, we present the simulation results of V2X spectrum sharing for V2I links and V2VRU links using the proposed SARL. We built our simulation using Python and TensorFlow software. For the evaluation methodology, we used the urban crossing environment described by 3GPP, which gives the vehicle mobility models, vehicle densities etc. [2]. We used previously reported channel characterization to implement the V2VRU communications [16]. We used the Rayleigh fast fading channel model and log-distance path loss model with log-normal shadowing distribution. The budget time constraint T used for V2VRU data packet transmissions is 100 ms, the time for path loss and shadowing update is 100 ms, and the time for fast fading update is 1 ms. The simulation parameters are listed in Table 1, which describes the channel models characterization for both V2I and V2VRU links. Table 2 describes the major simulation parameters, which are similar to those in [14]. For the V2VRU communication, where VRU

Table 2. Simulation Parameters Based on Reference [14]

Parameters	Values
Number of vehicles	4
Carrier Frequency	5.9 GHz
Bandwidth	10 MHz
RSU or BS height	25 m
RSU or BS antenna gain	8 dBi
RSU or BS receiver noise figure	5 dB
Vehicle or VRU antenna height	1.5 m
Vehicle or VRU antenna gain	3 dBi
Vehicle or VRU receiver noise figure	9 dBm
Vehicle speed	10 m/s
V2I transmit power	23 dBm
V2P transmit power	[23, 10, 5, -100] dBm
Noise power	-114 dB
V2P Packet size	[1..6] \times 1 KB
Time constraint for V2P packet transmission	100 ms
Path loss and shadowing update	Every 100 ms
Fast fading model	Rayleigh fading
Fast fading update	Every 1 ms

is a pedestrian, we used three scenarios based on [16]: (1) vehicle-to-static-VRU, (2) vehicle-to-moving-VRU, and (3) vehicle-to-VRU-with-crowd-shadowing.

The proposed DQN consists of 3 connected hidden layers with 500, 250, and 120 neurons. The DQN training lasts 3000 episodes. The packet size of VRU is 2 KB in the training phase, and it varies from 1 to 6 KB in the testing phase to verify the robustness of the proposed SARL algorithm.

We compared the proposed SARL for V2I and V2VRU spectrum sharing against the baseline of random resource allocation. The random baseline chooses the spectrum sub-band and transmission power in a random fashion at each time instant. We simulate 4 vehicles that are using 4 sub-bands where each sub-band has a bandwidth of 1 MHz. The transmission power for the V2VRU links is predetermined using the value of [23, 15, 5, -100] dBm, while the transmission power for V2I is fixed at 23 dBm. The value of -100 dBm is equivalent to 0 transmission power. We also compared the proposed solution against another baseline that used a fixed maximum transmission power for the V2P communication (23 dBm).

From Fig. 2, both the data rate and PRR decrease when the packet size increases due to the time budget constraint. The larger packet size will need a higher transmission power to improve the reception probability in the receiver side, hence leading to a higher interference. As we can see from Fig. 2, our proposed solution gave the best performance compared to the other baselines as it gave higher data rates for V2I communication. It also gave around the same performance for PRR of the Maximized V2P transmission power. However this maximized V2P transmission power gives a higher value of interference, hence reducing the data rate of V2I communication. In terms of performance metrics, the proposed SARL algorithm gives better performance compared to the random baseline. With a trained model using the packet size of 2 KB, we obtained 13% improvement in the data rate and a 4% improvement in PRR. Furthermore, the proposed SARL shows a consistently better performance with different packet

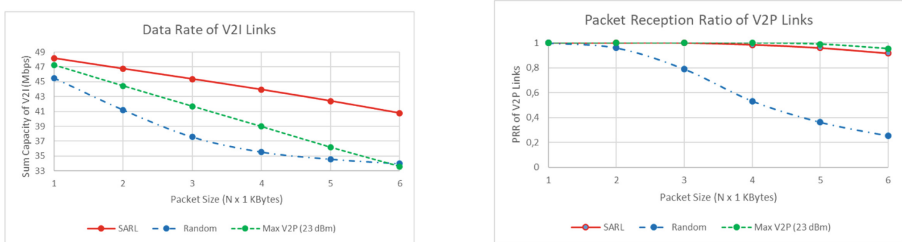


Fig. 2. The performance of SARL, random baseline and maximum transmission power for V2P (23 dBm) baseline in scenario 1 (Static Pedestrian).

sizes. For the packet size of 1 KB, the data rate improved by 6% even if the PRR did not improve significantly, while for the packet size of 6 KB, the data rate improvement was 20% and the PRR performance improved by around 67%.

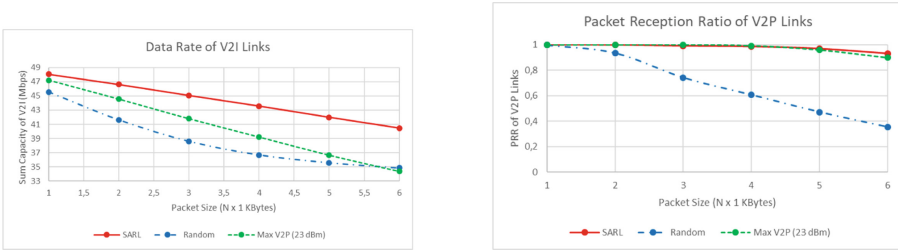


Fig. 3. The performance of SARL, random baseline and maximum transmission power for V2P (23 dBm) baseline in scenario 2 (moving pedestrian).

From Fig. 3, we can see that the proposed SARL offers a better performance for the scenario 2 which is spectrum sharing between V2I and V2VRU for the moving pedestrian. In the training phase, for the packet size of 2 KB, the SARL approach gives 12% improvement in data rate and 6% in PRR. For the smaller packet size of 1 KB, the performance of PRR is almost the same, and the performance of the data rate is improved by 5%; while for the larger packet size of 6 KB, the PRR improvement is 16% and the PRR performance is improved by 58%. The performance of the scenario 2 has the same results as in scenario 1, where the maximized transmission power of 23 dBm for the V2P communication gave around the same performance for the PRR of the proposed SARL solution. However, it reduced the data rate performance of the V2I communication due to the higher value of the interference.

Figure 4 shows the results for the scenario 3 which is the spectrum sharing between V2I and V2VRU for the pedestrian with crowd shadowing. In the training phase, for the packet size of 2KB, the SARL approach gives 16% improvement in data rate and 6% improvement of PRR. In the testing phase, for the smaller

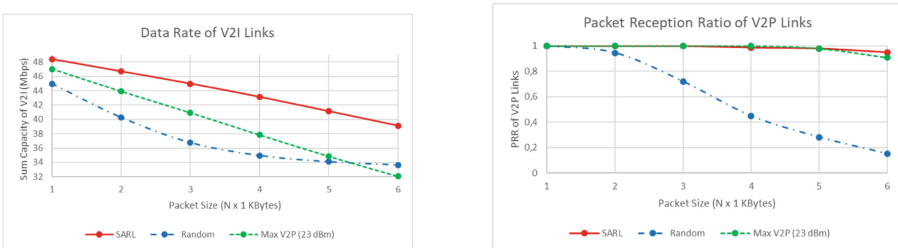


Fig. 4. The performance of SARL, random baseline and maximum transmission power for V2P (23 dBm) baseline in scenario 3 (crowded pedestrian).

packet size of 1 KB, the data rate improved by 8% while the performance of the PRR remained unchanged. While for the larger packet size of 6 KB, the data rate improvement was 16% and the PRR performance was improved by 80%. The data rate and PRR performances of the scenario 3 also gave the similar results to scenarios 1 and 2, where the proposed SARL solution gave the best performance and the maximized transmission power of the V2P communication gave around the same performance for the PRR of the V2P communication, but gave a reduced data rate for V2I communication due to high interference.

6 Conclusions

In this paper, we addressed the problem of optimizing the spectrum sharing of the V2X communication by improving the PRR of V2VRU links, and maximizing the data rate of V2I links. We proposed a DRL approach for spectrum sharing between V2I and V2VRU. A single-vehicle selects the spectrum sub-bands and the transmission power based on the trained DQN using acquired information from the environment. We compared our proposed model with a random baseline that randomly chooses the spectrum and power transmission. We also compared it with another baseline that uses the maximized transmission power of V2VRU communication (23 dBm). Our simulation result showed a significant improvement in the V2I data rate and V2VRU PRR for different VRU communication scenarios. The improvement of PRR implies more reliable V2VRU communication, which in turn improves the reliability of safety applications intended for VRU protection on the road. At the same time, the proposed SARL gave better data rates for V2I communications, thus enabling better performance for higher-data rate V2I-applications.

References

1. 3rd Generation Partnership Project (3GPP): Technical Specification Group Radio Access Network; Study on enhancement of 3GPP Support for 5G V2X Services; (Release 15). Technical Report (TR), Version 15.1.0, (2017)
2. 3rd Generation Partnership Project (3GPP): Technical Specification Group Radio Access Network; Study on LTE-based V2X Services; (Release 14). Technical Report (TR), Version 14.0.0, (2016)
3. Molina-Masegosa, R., Gozalvez, J.: LTE-V for Sidelink 5G V2X vehicular communications: a new 5G technology for short-range vehicle-to-everything communications. *IEEE Veh. Technol. Mag.* **12**(4), 30–39 (2017). <https://doi.org/10.1109/MVT.2017.2752798>
4. Alalewi, A., Dayoub, I., Cherkaoui, S.: On 5G–V2X use cases and enabling technologies: a comprehensive survey. *IEEE Access* **9**, 107710–107737 (2021). <https://doi.org/10.1109/ACCESS.2021.3100472>
5. Linget, T.: Vulnerable road user protection. 5GAA White Paper (2020)
6. Abid, M.A., Chakroun, O., Cherkaoui, S.: Pedestrian collision avoidance in vehicular networks. In: *IEEE International Conference on Communications (ICC 2013)*, pp. 2928–2932. IEEE, Budapest (2013). <https://doi.org/10.1109/ICC.2013.6654987>

7. Rezgui, J., et al.: Deterministic access for DSRC/802.11p vehicular safety communication. In: 7th International Wireless Communications and Mobile Computing Conference, pp. 595–600. IEEE, Istanbul (2011). <https://doi.org/10.1109/IWCMC.2011.5982600>
8. Azizian, M., et al.: An optimized flow allocation in vehicular cloud. *IEEE Access* **4**, 6766–6779 (2016). <https://doi.org/10.1109/ACCESS.2016.2615323>
9. Chakroun, O., et al.: Overhead-free congestion control and data dissemination for 802.11p VANETs. *Veh. Commun.* **1**(3), 123–133 (2014). <https://doi.org/10.1016/j.vehcom.2014.05.003>
10. Rezgui, J., et al.: About deterministic and non-deterministic vehicular communications over DSRC/802.11p. *Wireless Commun. Mobile Comput.* **14**(15), 1435–1449 (2014). <https://doi.org/10.1002/wcm.2270>
11. Azizian, M., et al.: Improved multi-channel operation for safety messages dissemination in vehicular networks. In: Proceedings of the fourth ACM International Symposium on Development and Analysis of Intelligent Vehicular Networks and Applications (DIVANet 2014), pp. 81–85. Association for Computing Machinery, New York (2014). <https://doi.org/10.1145/2656346.2656410>
12. Azizian, M., et al.: A distributed D-hop cluster formation for VANET. In: IEEE Wireless Communications and Networking Conference, pp. 1–6. IEEE, Doha (2016). <https://doi.org/10.1109/WCNC.2016.7564925>
13. Triwinarko, A., Cherkaoui, S., Dayoub, I.: Performance of radio access technologies for next generation V2VRU networks. In: IEEE International Conference on Communications (ICC 2022), pp. 1524–1529. IEEE, Seoul (2022). <https://doi.org/10.1109/ICC45855.2022.9838580>
14. Liang, L., Ye, H., Li, G.Y.: Spectrum sharing in vehicular networks based on multi-agent reinforcement learning. *IEEE J. Sel. Areas Commun.* **37**(10), 2282–2292 (2019). <https://doi.org/10.1109/JSAC.2019.2933962>
15. Mlika, Z., Cherkaoui, S.: Network slicing for vehicular communications: a multi-agent deep reinforcement learning approach. *Ann. Telecommun.* **76**(9), 665–683 (2021). <https://doi.org/10.1109/JSAC.2019.2933962>
16. Rashdan, I., de Ponte Muller, F., Wang, W., Schmidhammer, M., Sand, S.: Vehicle-to-Pedestrian channel characterization: wideband measurement campaign and first results. In: 12th European Conference on Antennas and Propagation (EuCAP 2018), 340 (5 pp.)–340 (5 pp.). Institution of Engineering and Technology (2018)
17. Mnih, V., et al.: Human-level control through deep reinforcement learning. *Nature* **518**(7540), 529–533 (2015)