Workshop Proceedings of the 19th International Conference on Intelligent Environments (IE2023)
G. Bekaroo et al. (Eds.)

doi:10.3233/AISE230018

Self-Adapted Resource Allocation in V2X Communication

Mansoor JAMAL [a], Zaib ULLAH [b] and Musarat ABBAS [a]
[a] *Quaid-I-Azam University, Islamabad, Pakistan*
[b] *Universita Telematica Giustino Fortunato, Benevento, Italy*

Abstract. The intelligent transportation system (ITS) along with vehicular communications are making our daily life safer and easier e.g., saving time, traffic control, safe driving, etc. Many transmission mode selection and resource allocation schemes are mitigating to full the quality of service requirements with minimum latency and negligible interference. For message transmission to far away vehicles realistic cellular link and mode selection is a bottleneck, however, for nearby devices, the safety critical information needs V2V links. Reinforcement learning (RL) and Deep learning (DL) have reshaped vehicular communication in a new model where vehicles act like human beings and take their decisions autonomously without human intervention. In our work, we investigate Vehicle-to-vehicle (V2V) and Vehicle-to-infrastructure where each link takes a decision to find the optimal subband and power level. We investigated the case where each V2V link connected with Vehicle-to-Infrastructure (V2I) satisfying stringent latency constraints while minimizing interference. We exploit RL with DL to develop highly intelligent performance where agents effectively learn to select mode and share spectrum with V2V and V2I.

Keywords. Intelligent transport system (ITS), Reinforcement learning (RL), Deep learning (DL), Vehicle-to-vehicle (V2V), Vehicle-to-infrastructure (V2I)

1. Introduction

Next-generation road vehicles are termed connected and automated vehicles (CAVs) due to their integrated IoT, sensors, and wireless communication technologies [1,2]. The sudden rise of smart driving perceptions has paid much attention to vehicular communication. The vehicular communication performance has improved due to its large-scale information sharing with other nearby vehicles and roadside communication units. There have been massive research endeavors by academia and the research industry to support various communication types, including multiple antenna communication [3,4], wireless sensor network [5], vehicle-to-vehicle (V2V), Vehicle-to-infrastructure (V2I), Vehicle-to-Network (V2N), and vehicle-to-pedestrian (V2P). The prior communication aspects between the vehicle and to rest of the communicating nodes are collectively known as Vehicle-to-Everything (V2X) communication. Consequently, V2X led to improvements in traffic efficiency, safety, time-saving, and the effective organization of modern transportation. The V2X communication supports infotainment, smart parking, lane changing, road crossing, speed limitation, and no parking, etc [6]. The challenge of communicating with the base station (BS) and other devices has gained sufficient research interest.

Both the organizations, the institute of electrical and electronics engineers (IEEE) and the third generation partnership project (3GPP) support device-to-device communication based on dedicated short-range communication and cellular network [7].
In V2X, two modes V2I and V2V are of high significance to instantaneously deliver critical information such as location, cooperative awareness messages, road conditions, etc . [8]. The abrupt rise in mobile technology raised the demand for a high data-rate cellular network to sustain video streaming, infotainment, and real-time V2X communication infrastructure [9]. In the critical scenario, a real-time communication link should be capable to transmit a data packet of 1200 bytes with the least latency of 5 ms, and with 100% reliability, [10].

Recently, with the development in computing technology and the introduction of new machine learning algorithms e.g. reinforcement learning [11], neural network [12] the goal of Artificial Intelligence (AI) has become a step closer. AI has important application in diverse fields including:healthcare [13,14,15], home medication [16], dynamic treatment regimes [17,18], UAVs, 5G and autonomous control [19] risk management [20,21], touristic planning [22] and communication [23].

Deep reinforcement learning (DRL) [24], is being investigated as a potential remedy for the slow convergence rate and suboptimal performance of established RL techniques when dealing with problems having a large state or action space. Deep reinforcement learning (DRL) is an advanced approach to combining deep learning with RL. RL is a subset of machine learning focused on directing policy. It supports decision-making by determining the incentives and policies utilized in decision-making to acquire the desired result. In addition, each agent in RL chooses the optimal action based on the learned policy [25]. By employing Deep Neural Networks (DNNs), DRL accelerates the learning procedure and enhances performance exceeding standard RL algorithms. Moreover, DRL effectively tackles complex network problems [26] and enables agents to select optimal strategies with little or no information interaction. As a result, DRL optimizes communication overhead and delay while improving reliability [24].

The rest of the article is organized as follows. Section II describes the most recent related works in a more abstract manner. Section III explains System Model and Problem Formulation. Section IV presents Deep Reinforcement Learning for Mode Selection and Resource Allocation. Section V talks about results and analysis, and Section VI concludes the paper.

2. Related Works

Research on D2D communication has been a fascinating area of study for the past many years. There are two modes for resource allocation, dedicated and shared resource allocation pool for cellular V2X communication [27]. In [28] based on deep reinforcement learning a decentralized mechanism of resource allocation has been proposed for cellular V2X communication. The sum capacity of V2I users has been maximized while guaranteeing V2V latency and reliability requirements. Based on deep reinforcement learning a novel mechanism for resource allocation has been proposed for V2V communication with small transmission overhead [29].
Traditional resource allocation mechanism for cellular-connected vehicular devices are

inapplicable due to fast channel variation and therefore requires full channel state information (CSI) rather than instantaneous. In [30] power allocation design and spectrum sharing have been investigated taking into account the required QOS of vehicular communication. To maximize system capacity and energy efficiency and minimize signal overhead for cellular V2X communication a DRL-based resource allocation mechanism has been adopted [31]. In [32] efficiency of the overall system, and throughput has been maximized in the V2X communication system by focusing on optimal transmit power allocation adopting a supervised learning technique for the deep neural network (DNN). Vehicle-to-Vehicle (V2) and vehicle-to-infrastructure (V2I) links shared resources have been investigated in the dynamic vehicular environments with effective and timely resource allocation policy [33].

With the increasing number of Internet of Things (IoT) devices, latency has always been a major constraint. In [34] this issue has been investigated through non-orthogonal multiple access (NOMA) considering user scheduling and power allocation for all users.

3. Technical Background

This section presents a brief introduction to RL. Machine learning has wide applications in image process, voice recognition, text analysis, object detection, etc.

As a branch of machine learning reinforcement learning focuses on independent decision-making and optimizing action policy by frequent interaction with the environment [35]. RL can be modeled as a Markov decision process to get an optimal policy for taking an action after observing an environment.

RL can be divided into two types based on transition probability and immediate reward, model-free and model-based learning, model-free algorithms like Q-learning are used in the stochastic environment when probability and reward functions are unknown. The optimal policy is updated using Q-table but when state space is too long then fewer Q values are updated after a long time and maximum iterations result in latency.

The structure of reinforcement learning is shown in Fig.1,

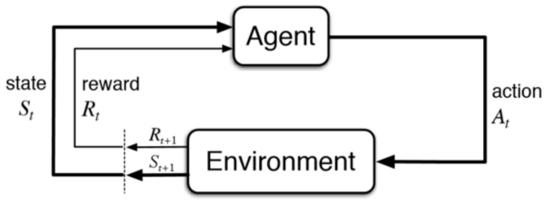

Figure 1. Deep RL for V2X Communication.

4. System Model and Problem Formulation

In our model vehicular communication consists of two types of vehicular user equipment (VUE). $\mathcal{N} = \{1,2,...,N\}$ cellular user pair (CUP) and $\mathcal{M} = \{1,2,...,M\}$ vehicular user pair (VUP). CUP is demanding realistic Vehicle-to-infrastructure (V2I) links via the cellular network for communication with a base station (BS) to share safety critical messages, road conditions, location, etc. As shown in Fig.2.

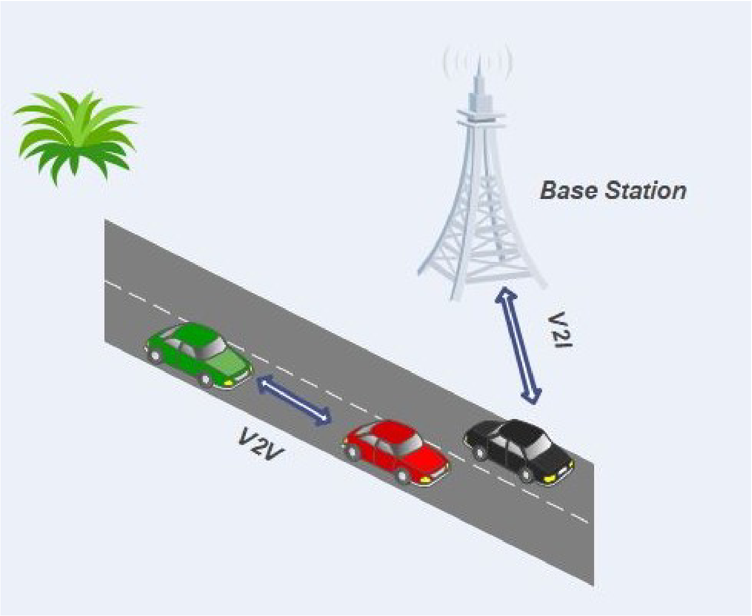

Figure 2. An illustrative structure of V2X Communication

The spectrum is orthogonally allocated to CUP. VUP needs Vehicle-to-vehicle (V2V) links to share cooperative awareness messages (CAM) for traffic and safety management with the other VUP consists of one transmitter and one receiver. To improve spectrum utilization efficiency and to reduce interference at the base station the orthogonally allocated to CUP up-link spectrum can be reused by VUP.
The interference to the CUP consists of two parts: the background noise and the VUP sharing the same resource block. The SINR for the ınth CUP will be

$$\gamma_n^c = \frac{P_n^c h_n}{\sigma^2 + \sum_{\imath m \varepsilon \mathscr{M}} r_{m[n]} P_m^v g_m} \tag{1}$$

Where P_n^c and P_m^v indicate transmit power of CUP n^{th} and VUP m^{th} , σ^2 is the noise power, h_n is the power gain of the channel corresponding to the n^{th} CUP, g_m is the interference power gain of the m^{th} VUP, $r_{m[n]}$ is the indicator function for spectrum allocation with $r_{m[n]} = 1$ if the m^{th} VUP reuses the spectrum of n^{th} CUP and $r_{m[n]} = 0$, otherwise. Hence the achievable data rate of n^{th} CUP with bandwidth W can be written as

$$R_n^c = W.log_2(1 + \gamma_n^c) \tag{2}$$

Similarly, each VUP can switch their selected mode whether to communicate with other VUP directly or with CUP indirect communication through BS. The communica-

tion mode selection of VUP can be shown as $s_m = 1$ if VUP selects the CUP mode, otherwise selects V2V (Vehicle-to-Vehicle) mode.
As VUP can communicate directly with VUP and can also with CUP indirectly so the interference comes from CUP and VUP sharing the same resource block (RB). The SINR of the $m_t h$ VUP can be expresses as

$$\gamma_m^v = \frac{P_m^v h_m}{\sigma^2 + H_c + H_v} \tag{3}$$

with

$$H_c = \sum_{\imath n \varepsilon \mathscr{N}} P_n^c g_n r_{m[n]} s_m \tag{4}$$

and

$$H_v = \sum_{\imath j \varepsilon \mathscr{M}} \sum_{\imath m \varepsilon \mathscr{M}} r_{m[n]} r_{j[n]} P_j^v g_j \tag{5}$$

Where h_m is the power gain of the channel corresponding to m^{th} VUP, g_n is the interference power gain of n^{th} CUP, where ıj is for the VUP sharing the same resource block.Hence the achievable data rate of the m^{th} VUP can be written as

$$R_m^v = W.log_2(1 + \gamma_m^v) \tag{6}$$

Where an agent observes the environment, everything outside of a selected vehicular user pair and cellular user pair is considered part of the environment and their action such as spectrum selection and mode selection are part of environmental noise. At each time t VUE pair observes the environment as a state s_t, from state space, $\mathscr{S}$, and takes action a_t, from action space, $\mathscr{A}$, corresponds to how to select mode, resources, and power level for transmission following the policy π. When an agent takes an action the environment changes to a new state, s_{t+1}, and based on that action each agent receives reward r_t determine by the channel data rate of CUE and VUE.
The state observable to each agent consists of several state parts: the instantaneous channel information of the corresponding V2V links, G_t, channel gain from VUP transmitter to its corresponding VUP receiver and BS, $H_{N,t}$,$H_{M,t}$, received interference power on BS at previous sub-frame, $I_{t-1} = (I_{t-1}[1], ..., I_{t-1}[K])$, number of selected neighbors on each RB, $O_{t-1} = (O_{t-1}[1], ..., O_{t-1}[K])$, current load and remaining time to meet latency threshold, L_t, T_t. Thus the state space can be expressed as

$$s_t = \{G_t, H_{N,t}, H_{M,t}, I_{t-1}, O_{t-1}, L_t, T_t\}.$$

At each time the agent observes the state s_t from the state space $\mathscr{S}$ following the policy π, takes an action a_t from the action space $\mathscr{A}$ which consist of selecting a sub-channel, mode selection and power level for transmission. An agent can select the mode for transmission with a frequency band and power level that has small interference to all V2I and V2V links without violating latency constraints.

5. Results

In this section, the performance of the proposed method for cellular communications is evaluated through simulations. For the simulation, we consider a system with a carrier frequency of 2 GHz, where vehicles are dropped randomly according to the spatial Poisson process where each vehicle as an agent is able to communicate with three nearby vehicles. Our deep Q-network is fully connected and constructed by an input layer, a hidden layer, and an output layer containing 500, 250, 120 neurons in the three hidden layers. The learning rate at the beginning is 0.01 and decreases exponentially.

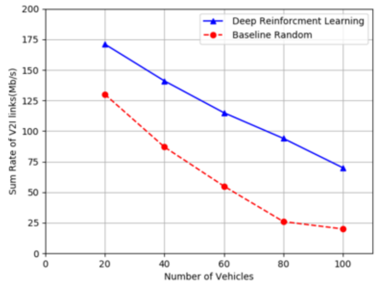

Figure 3. Sum Rate of V2I vs Number of Vehicles.

5.1. V2I Capacity

Fig. 3 shows the sum rate of V2I vs the number of vehicles. From the figure, we can infer that, with the increase in the number of vehicles, the number of V2V links increases as a result, the interference with the V2I link grows, and therefore the V2I capacity will drop.

5.2. V2V Latency

Fig.4 shows the probability that the V2V links satisfy the latency constraint versus the number of vehicles. From the figure, we can infer that, with the increase in the number of vehicles, the V2V links increase, as a result, it is more difficult to ensure every vehicle satisfies the latency constraint.

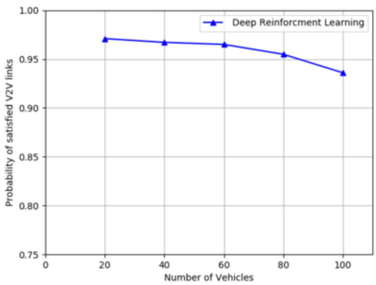

Figure 4. Probability of Satisfied V2V links vs the number of vehicles.

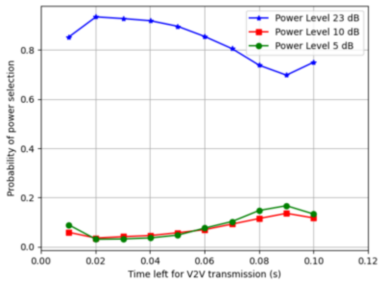

Figure 5. The Probability of power level selection with the remaining time for transmission.

5.3. *Power Level selection*

Fig.5 shows the probability for the agent to choose power levels with different time left for transmission when Double-Deep Q-Learning is used. The probability for the agent to choose the maximum power is decreased compared to the figure-3 when there is abundant time for transmission. Also, the probability of selecting maximum power to ensure the V2V latency constraint when a small amount of time left is increased. Apart from this, when the agent has abundant time for transmission it will select low-power transmission to reduce resource usage.

6. Conclusion

In this paper, a deep reinforcement learning-based decentralized mechanism has been proposed for vehicular communications. From the simulation results, it can be observed that each agent can intelligently learn how to satisfy V2V constraints in terms of latency, power, and transmission overhead while minimizing interference to V2I links.

References

[1] Faran Awais Butt, Jawwad Nasar Chattha, Jameel Ahmad, Muhammad Umer Zia, Muhammad Rizwan, and Ijaz Haider Naqvi. On the integration of enabling wireless technologies and sensor fusion for next-generation connected and autonomous vehicles. *IEEE Access*, 10:14643–14668, 2022.

[2] M Nadeem Ahangar, Qasim Z Ahmed, Fahd A Khan, and Maryam Hafeez. A survey of autonomous vehicles: Enabling communication technologies and challenges. *Sensors*, 21(3):706, 2021.

[3] Muddasar Naeem, Sajid Bashir, Zaib Ullah, and Aqeel A Syed. A near-optimal scheduling algorithm for efficient radio resource management in multi-user mimo systems. *Wireless Personal Communications*, 106(3):1411–1427, 2019.

[4] Muddasar Naeem, Sajid Bashir, Muhammad Usman Khan, and Aqeel A Syed. Performance comparison of scheduling algorithms for mu-mimo systems. In *2016 13th International Bhurban Conference on Applied Sciences and Technology (IBCAST)*, pages 601–606. IEEE, 2016.

[5] Alessandro Testa, Antonio Coronato, Marcello Cinque, and Juan Carlos Augusto. Static verification of wireless sensor networks with formal methods. In *2012 Eighth International Conference on Signal Image Technology and Internet Based Systems*, pages 587–594. IEEE, 2012.

[6] Sohan Gyawali, Shengjie Xu, Yi Qian, and Rose Qingyang Hu. Challenges and solutions for cellular based v2x communications. *IEEE Communications Surveys & Tutorials*, 23(1):222–255, 2020.

[7] Khadige Abboud, Hassan Aboubakr Omar, and Weihua Zhuang. Interworking of dsrc and cellular network technologies for v2x communications: A survey. *IEEE Transactions on Vehicular Technology*, 65(12):9457–9470, 2016.
[8] Ning Lu, Nan Cheng, Ning Zhang, Xuemin Shen, and Jon W Mark. Connected vehicles: Solutions and challenges. *IEEE internet of things journal*, 1(4):289–299, 2014.
[9] Peng Liu, Chaoyu Wang, Tingting Fu, and Yue Ding. Exploiting opportunistic coding in throwbox-based multicast in vehicular delay tolerant networks. *IEEE Access*, 7:48459–48469, 2019.
[10] Trung-Kien Le, Umer Salim, and Florian Kaltenberger. An overview of physical layer design for ultra-reliable low-latency communications in 3gpp releases 15, 16, and 17. *IEEE access*, 9:433–444, 2020.
[11] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT press, 2018.
[12] Muddasar Naeem, Giovanni Paragiola, Antonio Coronato, and Giuseppe De Pietro. A cnn based monitoring system to minimize medication errors during treatment process at home. In *Proceedings of the 3rd International Conference on Applications of Intelligent Systems*, pages 1–5, 2020.
[13] Antonio Coronato, Muddasar Naeem, Giuseppe De Pietro, and Giovanni Paragliola. Reinforcement learning for intelligent healthcare applications: A survey. *Artificial Intelligence in Medicine*, 109:101964, 2020.
[14] Antonio Coronato and Muddasar Naeem. A reinforcement learning based intelligent system for the healthcare treatment assistance of patients with disabilities. In *Pervasive Systems, Algorithms and Networks: 16th International Symposium, I-SPAN 2019, Naples, Italy, September 16-20, 2019, Proceedings*, pages 15–28. Springer, 2019.
[15] Muddasar Naeem and Antonio Coronato. An ai-empowered home-infrastructure to minimize medication errors. *Journal of Sensor and Actuator Networks*, 11(1):13, 2022.
[16] Mario Ciampi, Antonio Coronato, Muddasar Naeem, and Stefano Silvestri. An intelligent environment for preventing medication errors in home treatment. *Expert Systems with Applications*, 193:116434, 2022.
[17] Syed Ihtesham Hussain Shah, Antonio Coronato, Muddasar Naeem, and Giuseppe De Pietro. Learning and assessing optimal dynamic treatment regimes through cooperative imitation learning. *IEEE Access*, 10:78148–78158, 2022.
[18] Muddasar Naeem, Antonio Coronato, and Giovanni Paragliola. Adaptive treatment assisting system for patients using machine learning. In *2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS)*, pages 460–465. IEEE, 2019.
[19] Zaib Ullah, Fadi Al-Turjman, Uzair Moatasim, Leonardo Mostarda, and Roberto Gagliardi. Uavs joint optimization problems and machine learning to improve the 5g and beyond communication. *Computer Networks*, 182:107478, 2020.
[20] Giovanni Paragliola, Antonio Coronato, Muddasar Naeem, and Giuseppe De Pietro. A reinforcement learning-based approach for the risk management of e-health environments: A case study. In *2018 14th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*, pages 711–716. IEEE, 2018.
[21] Syed Ihtesham Hussain Shah, Muddasar Naeem, Giovanni Paragliola, Antonio Coronato, and Mykola Pechenizkiy. An ai-empowered infrastructure for risk prevention during medical examination. *Expert Systems with Applications*, page 120048, 2023.
[22] Antonio Coronato, Claudia Di Napoli, Giovanni Paragliola, and Luca Serino. Intelligent planning of onshore touristic itineraries for cruise passengers in a smart city. In *2021 17th International Conference on Intelligent Environments (IE)*, pages 1–7. IEEE, 2021.
[23] Muddasar Naeem, Giuseppe De Pietro, and Antonio Coronato. Application of reinforcement learning and deep learning in multiple-input and multiple-output (mimo) systems. *Sensors*, 22(1):309, 2021.
[24] Ishan Budhiraja, Neeraj Kumar, and Sudhanshu Tyagi. Deep-reinforcement-learning-based proportional fair scheduling control scheme for underlay d2d communication. *IEEE Internet of Things Journal*, 8(5):3143–3156, 2020.
[25] Shalini Yadav and Rahul Rishi. Joint mode selection and resource allocation for cellular v2x communication using distributed deep reinforcement learning under 5g and beyond networks. *Available at SSRN 4356711*.
[26] Muddasar Naeem, Antonio Coronato, Zaib Ullah, Sajid Bashir, and Giovanni Paragliola. Optimal user scheduling in multi antenna system using multi agent reinforcement learning. *Sensors*, 22(21):8278, 2022.
[27] Mario H. Castañeda Garcia, Alejandro Molina-Galan, Mate Boban, Javier Gozalvez, Baldomero Coll-

Perales, Taylan Şahin, and Apostolos Kousaridas. A tutorial on 5g nr v2x communications. *IEEE Communications Surveys Tutorials*, 23(3):1972–2026, 2021.

[28] Xinran Zhang, Mugen Peng, Shi Yan, and Yaohua Sun. Deep-reinforcement-learning-based mode selection and resource allocation for cellular v2x communications. *IEEE Internet of Things Journal*, 7(7):6380–6391, 2020.

[29] Hao Ye, Geoffrey Ye Li, and Biing-Hwang Fred Juang. Deep reinforcement learning based resource allocation for v2v communications. *IEEE Transactions on Vehicular Technology*, 68(4):3163–3173, 2019.

[30] Le Liang, Geoffrey Ye Li, and Wei Xu. Resource allocation for d2d-enabled vehicular communications. *IEEE Transactions on Communications*, 65(7):3186–3197, 2017.

[31] Hyebin Park and Yujin Lim. Deep reinforcement learning based resource allocation with radio remote head grouping and vehicle clustering in 5g vehicular networks. *Electronics*, 10(23):3015, 2021.

[32] Jin Gao, Muhammad R. A. Khandaker, Faisal Tariq, Kai-Kit Wong, and Risala T. Khan. Deep neural network based resource allocation for v2x communications. In *2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall)*, pages 1–5, 2019.

[33] Yi Yuan, Gan Zheng, Kai-Kit Wong, and Khaled B Letaief. Meta-reinforcement learning based resource allocation for dynamic v2x communications. *IEEE Transactions on Vehicular Technology*, 70(9):8964–8977, 2021.

[34] Huiyi Ding and Ka-Cheong Leung. Resource allocation for low-latency noma-v2x networks using reinforcement learning. In *IEEE INFOCOM 2021 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pages 1–6, 2021.

[35] Muddasar Naeem, Syed Tahir Hussain Rizvi, and Antonio Coronato. A gentle introduction to reinforcement learning and its application in different fields. *IEEE access*, 8:209320–209344, 2020.