A Learning-based Resource Slicing Algorithm to Achieve URLLC in Vehicle-to-Vehicle Networks

Madyan Alsenwi, Shashi Raj Pandey, and Choong Seon Hong
Department of Computer Science and Engineering, Kyung Hee University, South Korea
malsenwi@khu.ac.kr, shashiraj@khu.ac.kr, cshong@khu.ac.kr

Abstract

Vehicle-to-Vehicle (V2V) communications are mainly used in exchanging data safety driving information among the neighboring vehicles which requires Ultra-Reliable and Low-Latency Communication (URLLC). In view of this, an efficient resource slicing mechanism is required to allocate resources to V2V links considering the required transmission reliability and latency. Recently, many works have considered the shared resources approach where each V2V link decides the allocated resources in a decentralized manner to improve the spectrum efficiency. However, sharing resources among vehicles causes interference which may violate the required transmission reliability. To this end, in this paper, we propose a resource slicing algorithm in which the next Generation Node B (gNB) allocates orthogonal Resource Blocks (RBs) to each V2V link to mitigate the interference caused by sharing resources and hence improve the transmission reliability. Considering the dynamic nature of V2V networks and to improve the spectrum efficiency, the gNB updates the allocated resources to each V2V link periodically. To achieve that, we leverage the capability of the reinforcement learning (RL) algorithms in doing decisions by interacting with the environment to design a resource slicing algorithm that considers the channel state of each V2V link. Simulation results show the performance of the proposed algorithm.

Keywords - V2V communication, URLLC, resource slicing, reinforcement learning.

1. Introduction

Recently, vehicle-to-vehicle (V2V) communications have attracted attention as a key enabler of intelligent transportation systems. In this regard, reliable and low latency communication is required to allow efficient cooperation among vehicles. Many applications could be supported in vehicular networks such as information dissemination, internet access, mobility enhancement, and accident alarming [1, 2]. Recent advances in cellular networks have been considered a promising solution to achieve high reliability, low latency communications in highly-dynamic environments. In the fifth-generation (5G) wireless networks, vehicles are allowed to communicate with each other directly through the licensed cellular spectrum which is controlled by the gNB. Compared to the IEEE 802.11p based communication, which is a contention-based communication, cellular-based V2V communication has a higher data rate, lower latency, and higher reliability [3, 4].

Works on V2V communication have gained attention in both academia and industry in recent years. For instance, the authors in [5] derived the lower bound of the tradeoff among the spectrum efficiency, reliability, and latency. They formulated an optimization problem to find the optimal frame design aiming at minimizing the transmission latency. The work in [6] considered the joint communication mode selection and resource allocation in V2V based internet of vehicle networks. The objective is to satisfy the required reliability and latency of V2V links while maximizing the sum data rate of the internet of vehicles network. Most recent studies on V2V communications have considered that the frequency channels are shared among vehicles to improve spectrum efficiency. However, this degrades the transmission reliability due to the interference among V2V links.

Motivated by the aforementioned facts, we propose an efficient resource slicing framework in which the gNB allocates dedicated resources to each V2V link based on the traffic and channel state of each V2V link to ensure the required transmission reliability. To improve the spectrum efficiency, the allocated resources are updated dynamically at each time slot based on the current network state. In doing so, we propose a reinforcement learning-based algorithm that can interact with the network environment and do decisions based on the current network state. In the proposed algorithm, the reward function is designed to consider the required transmission reliability of each V2V link. Furthermore, leveraging the 5G New Radion (NR) numerologies, the generated packets at each vehicle are immediately transmitted over short transmission time intervals (sTTI) to satisfy the required latency.

2. System Model

We consider a wireless network, where a number of vehicles communicate with each other under the coverage of a gNB as shown in Fg. 1. Let $\mathcal{N}$, and $\mathcal{K}$ denote the set of V2V links, and the set of the available RBs at the gNB for the V2V communications, respectively. Each vehicle transmits short packets size with a fixed length of μ. RBs are allocated to each V2V link by the gNB. We consider the Finite Block Length (FBL) in calculating the data rate of each V2V link due to the short length of the generated packets at each vehi-

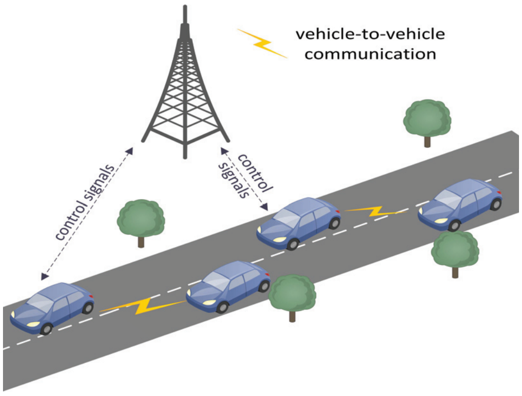

Figure 1: System model.

cle. Therefore, the achievable data rate of the n^{th} V2V link in the FBL regime is approximated as [7, 8]

$$r_k = \frac{f|\mathcal{N}_k|}{\ln 2}\left[\ln\left(1+\gamma_k\right) - \sqrt{\frac{D_k}{\tau f|\mathcal{N}_k|}}Q^{-1}(\epsilon_k)\right], \quad (1)$$

$$\gamma_k = \frac{p_k|h_k|^2}{\sigma^2}, \qquad D_k = 1 - \frac{1}{\left(1+\gamma_k\right)^2},$$

where $\mathcal{N}_k$ is the set of RBs assigned to the k^{th} V2V link and $|\mathcal{N}_k|$ is the cardinality; $Q^{-1}(\cdot)$ is the inverse of the Gaussian Q-function, f is the RB bandwidth, h_k is the channel gain, p_k is the transmission power, and ϵ_k is the transmission error probability,

We define the transmission reliability of each V2V link in terms of the probability of transmission error ϵ_k of each V2V link. Therefore, resources are allocated to each V2V link such that

$$\epsilon_k \leq \epsilon_k^{\max},\ \forall\ k \in \mathcal{K}, \quad (2)$$

where $\epsilon_k^{\max}$ is the maximum allowed transmission error probability of the k^{th} V2V link. The data rate of the k^{th} V2V link in (1) can be expressed as

$$\frac{\mu_k \times l_k}{\tau} \approx \frac{f|\mathcal{N}_k|}{\ln 2}\left[\ln(1+\gamma_k) - \sqrt{\frac{D_k}{\tau f|\mathcal{N}_n|}}Q^{-1}(\epsilon_k)\right], \quad (3)$$

where τ is the transmission time which equals the sTTI length, and $l_k(t)$ is the number the number of generated packets at the k^{th} V2V link. Accordingly, the reliability constraint in (2) can be expressed as

$$Q\left(\frac{\ln(1+\gamma_k) - \frac{\mu_k \ln 2}{\tau f|\mathcal{N}_k|}}{\sqrt{\frac{D_k}{\tau f|\mathcal{N}_k|}}}\right) \leq \epsilon_k^{\max}. \quad (4)$$

3. Proposed RL-Based Resource Slicing Algorithm

In this section, we propose a RL-based algorithm to obtain a dynamic and real-time resource slicing for V2V communications. In particular, RL algorithms can make decisions in real-time by interacting with the environment making it an appropriate choice to work along in addressing resource allocation issues and decision making under uncertainty. In general, a RL algorithms are defined by the action space $\mathcal{A}$, state space $\mathcal{S}$, and reward $R(t)$. A RL algorithm can take an action $\boldsymbol{a}(t) \in \mathcal{A}$ given a state $\boldsymbol{s}(t) \in \mathcal{S}$ and gets the reward $R(t)$.

State space: we consider the state space as the state of each V2V link, i.e., number of generated packets and channel variations, at each time slot (decision epoch). Therefore, at time slot t, the state is given as $\boldsymbol{s}(t) = \{l_k(t), h_k(t),\ \forall k \in \mathcal{K}\}$.

Action space: we define $\boldsymbol{z}$ as an $K \times N$ binary indicator matrix, such that $z_{k,n}$ indicates whether RB n is assigned to V2V link k. Thus, the action space can be defined $\boldsymbol{a}(t) = \{z_{k,n}(t),\ \forall k \in \mathcal{K},\ n \in \mathcal{N}\}$.

Reward: considering the required transmission reliability of each V2V link, we define the reward function as

$$R(t) = \sum_{k\in\mathcal{K}}\sum_{N\in\mathcal{N}}\left[\epsilon_k^{\max} - Q\left(\frac{\ln(1+\gamma_k(t)) - \frac{\mu_k l_k \ln 2}{\tau f z_{k,n}(t)}}{\sqrt{\frac{D_k}{\tau f z_{k,n}(t)}}}\right)\right]. \quad (5)$$

The agent (gNB) chooses a policy $\pi(\boldsymbol{a}, \boldsymbol{s}) = \{\pi_k^n,\ \forall k \in \mathcal{K}, n \in \mathcal{N}\}$, where π_k^n is the probability of allocating the RB k to v2v link k. Specifically, the gNB observes the network state $\boldsymbol{s}(t)$ and selects an action based on the policy. Next, the reward $R(t)$ is calculated from (5). Finally, a new policy is learned in the next time slot based on the feedback from the environment.

Let $Q^{\pi}(\boldsymbol{s}, \boldsymbol{a})$ denote the Q-function which measures the quality of an action $\boldsymbol{a}(t)$ at a given state $\boldsymbol{s}(t)$, defined as

$$Q^{\pi}(\boldsymbol{s}, \boldsymbol{a}) = \mathbb{E}\left[\sum_{t=0}^{\infty}\gamma(t)R\big(\boldsymbol{s}(t), \boldsymbol{a}(t)\big)|s_0 = s, \pi\right]. \quad (6)$$

The $Q^{\pi}(\boldsymbol{s}, \boldsymbol{a})$ function is calculated using the Bellman equation [9, 10]:

$$Q^{\pi}(\boldsymbol{s}, \boldsymbol{a}) = \mathbb{E}\left[R\big(\boldsymbol{s}(t), \boldsymbol{a}(t)\big) + Q^{\pi}\big(\boldsymbol{s}(t+1), \boldsymbol{a}(t+1)\big)\right]. \quad (7)$$

The Q-function can be iteratively estimated using the Bellman equation as

$$Q^{t+1}(\boldsymbol{s}, \boldsymbol{a}) = \mathbb{E}\Big[R(t) + \max_{a(t+1)} Q^{t}(\boldsymbol{s}(t+1), \boldsymbol{a}(t))|\boldsymbol{s}(t), \boldsymbol{a}(t)\Big]. \quad (8)$$

The ϵ-greedy exploration method is considered to allow the agent to randomly select an action with a probability ϵ. In this method, the agent chooses the current estimated action with probability $1-\epsilon$ [9].

4. Performance Evaluation

In this section, we evaluate the performance of the proposed resource slicing algorithm. Several vehicles that communi-

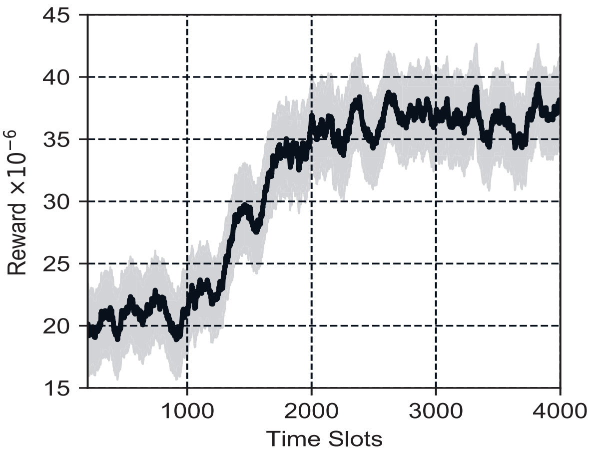

Figure 2: Convergence of the reward function.

cate with each other are considered under the coverage of a gNB. Each RB is considered to be composed of 12 subcarriers with a subcarrier-spacing of 15 kHz. The total available bandwidth at the gNB for V2V communications is considered to be 20 MHz. We consider the arrival of URLLC packets at each vehicle follows the Poisson process. The size of each URLLC packet is set to 32 bytes.

We study the convergence of the proposed RL-based resource slicing algorithm in Fig. 2. Specifically, Fig. 2 plots the reward function in (5) with time slots. As shown in Fig. 2, the values of the reward function increase with time, and this means that the difference between the threshold transmission error and the achieved transmission error increases. Thus, the proposed algorithm can efficiently reduce the transmission error to lower than the predefined threshold. Hence, the proposed resource slicing algorithm satisfies the required transmission reliability by V2V communications.

5. Conlusion

In this paper, we have studied the resource slicing problem in V2V communication networks. In particular, we have proposed a RL-based algorithm that allows the gNB to allocate orthogonal resources to each V2V link to ensure the required transmission reliability. Moreover, the required latency is ensured by transmitting the generated packets at each vehicle immediately over sTTI. We have provided simulation results to validate the efficiency of the proposed algorithm in satisfying the required transmission reliability. In future work, we will consider the coexistence problem of eMBB users and V2V communications. A deep RL-based algorithm would be considered to solve such a complicated problem.

Acknowledgement

This work was supported by the Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No.2019-0-01287, Evolvable Deep Learning Model Generation Platform for Edge Computing). *Dr. CS Hong is the corresponding author.

References

[1] M. Alsenwi, N. H. Tran, M. Bennis, A. K. Bairagi, and C. S. Hong, "eMBB-URLLC resource slicing: A risk-sensitive approach," *IEEE Communications Letters*, vol. 23, no. 4, pp. 740–743, apr 2019.

[2] F. Abbas, P. Fan, and Z. Khan, "A novel low-latency v2v resource allocation scheme based on cellular v2x communications," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 6, pp. 2185–2197, jun 2019.

[3] M. Alsenwi, I. Yaqoob, S. R. Pandey, Y. K. Tun, A. K. Bairagi, L. won Kim, and C. S. Hong, "Towards coexistence of cellular and WiFi networks in unlicensed spectrum: A neural networks based approach," *IEEE Access*, vol. 7, pp. 110 023–110 034, 2019.

[4] M. Alsenwi, H. Elsayed, and M. Elghoneimy, "Optimization of channel selection in cognitive heterogeneous wireless networks," in *2015 11th International Computer Engineering Conference (ICENCO)*. IEEE, dec 2015.

[5] H. Yang, K. Zhang, K. Zheng, and Y. Qian, "Joint frame design and resource allocation for ultra-reliable and low-latency vehicular networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 5, pp. 3607–3622, may 2020.

[6] H. Yang, X. Xie, and M. Kadoch, "Intelligent resource management based on reinforcement learning for ultra-reliable and low-latency IoV communication networks," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 5, pp. 4157–4169, may 2019.

[7] M. Alsenwi, S. R. Pandey, Y. K. Tun, K. T. Kim, and C. S. Hong, "A chance constrained based formulation for dynamic multiplexing of eMBB-URLLC traffics in 5g new radio," in *2019 International Conference on Information Networking (ICOIN)*. IEEE, jan 2019.

[8] M. Alsenwi, Y. K. Tun, S. R. Pandey, N. N. Ei, and C. S. Hong, "UAV-assisted multi-access edge computing system: An energy-efficient resource management framework," in *2020 International Conference on Information Networking (ICOIN)*. IEEE, jan 2020.

[9] R. Li, Z. Zhao, Q. Sun, C.-L. I, C. Yang, X. Chen, M. Zhao, and H. Zhang, "Deep reinforcement learning for resource management in network slicing," *IEEE Access*, vol. 6, pp. 74 429–74 441, 2018.

[10] M. Alsenwi, N. H. Tran, M. Bennis, S. R. Pandey, A. K. Bairagi, and C. S. Hong, "Intelligent resource slicing for embb and urllc coexistence in 5g and beyond: A deep reinforcement learning based approach," *arXiv preprint arXiv:2003.07651 (2020)*.