



## Article

# Reinforcement Learning Approach for Adaptive C-V2X Resource Management

Teguh Indra Bayu<sup>1,2</sup>, Yung-Fa Huang<sup>3,\*</sup> and Jeang-Kuo Chen<sup>1</sup>

<sup>1</sup> Department of Information Management, Chaoyang University of Technology, Taichung 413310, Taiwan; s10814906@gm.cyut.edu.tw or teguh.bayu@uksw.edu (T.I.B.); jkchen@cyut.edu.tw (J.-K.C.)

<sup>2</sup> Department of Informatics Engineering, Satya Wacana Christian University, Salatiga 50711, Indonesia

<sup>3</sup> Department of Information and Communication Engineering, Chaoyang University of Technology, Taichung 413310, Taiwan

\* Correspondence: yfahuang@cyut.edu.tw

**Abstract:** The modulation coding scheme (MCS) index is the essential configuration parameter in cellular vehicle-to-everything (C-V2X) communication. As referenced by the 3rd Generation Partnership Project (3GPP), the MCS index will dictate the transport block size (TBS) index, which will affect the size of transport blocks and the number of physical resource blocks. These numbers are crucial in the C-V2X resource management since it is also bound to the transmission power used in the system. To the authors' knowledge, this particular area of research has not been previously investigated. Ultimately, this research establishes the fundamental principles for future studies seeking to use the MCS adaptability in many contexts. In this work, we proposed the application of the reinforcement learning (RL) algorithm, as we used the Q-learning approach to adaptively change the MCS index according to the current environmental states. The simulation results showed that our proposed RL approach outperformed the static MCS index and was able to attain stability in a short number of events.

**Keywords:** MCS; reinforcement learning; Q-learning



**Citation:** Bayu, T.I.; Huang, Y.-F.; Chen, J.-K. Reinforcement Learning Approach for Adaptive C-V2X Resource Management. *Future Internet* **2023**, *15*, 339. <https://doi.org/10.3390/fi15100339>

Academic Editor: Ping Wang

Received: 28 August 2023

Revised: 28 September 2023

Accepted: 14 October 2023

Published: 15 October 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The introduction of vehicular communication technology is poised to significantly transform the transportation and mobility industry, enabling enhanced safety measures and increased efficiency in traffic management. Moreover, the advancement of vehicle autonomy, facilitated by improved sensing capabilities, is anticipated to bring out novel advanced services such as cooperative sensing and maneuvers. These services will heavily rely on direct sidelink connection. Vehicle communications criteria have primarily relied on IEEE 802.11p technology in the last ten years. Recently, the 3GPP has expeditiously formulated cellular standards to tackle the obstacles associated with vehicular communications, collectively called C-V2X [1]. This endeavor commenced with Release 14, also recognized as LTE-V, and has since been succeeded by Release 15 and 16, commonly denoted as NR-V2X [2].

Vehicular communications may be facilitated via two main approaches: using the existing cellular network infrastructure, explicitly employing C-V2X Mode 3 or NR-V2X Mode 1, or enabling direct communication between cars in an ad-hoc manner, known as C-V2X Mode 4 or NR-V2X Mode 2 [3]. Within the context of vehicular sidelink communication, cars possess the capability to independently assign and oversee a predetermined quantity of resources for specific time intervals. Radio resource management is accomplished via the implementation of a scheduling technique referred to as sensing-based semi-persistent scheduling (SB-SPS). Although resources are allocated semi-persistent, sicvehicles can adjust characteristics such as transmission power or MCS. Hence, vehicles can adjust these

characteristics to serve various objectives, such as managing traffic congestion or facilitating the transfer of packets of varying sizes [4].

In contrast to the uncomplicated process of adjusting transmission power during current transmissions in the C-V2X sidelink, modifying the MCS presents a more complex challenge due to its potential impact on resource occupancy. This challenge occurs because various MCS configurations alter the quantity of transmission resources, affecting the transmission bandwidth. The factors mentioned above have further implications for calculating power levels in the channel and the functioning of SB-SPS. The latter depends on distinct power measures to determine and allocate the necessary resources.

The link adaptation (LA) technique has recently been an interesting discussion in wireless communication technology research and the adaptive modulation and coding scheme (AMC) mechanism [5]. The involvement of machine learning is unavoidable in this matter. For instance, the utilization of a deep convolutional neural network (DCNN) is reported in [6]. The estimated combined channel matrix and the noise standard deviation are used as input features to train the DCNN. In [7], convolutional neural networks (CNNs) are proposed for MCS selection at the time of transmission in time-division-duplex (TDD) systems. The proposed method determines the most suitable modulation and coding scheme (MCS) for future transmission based on the shifts in received signal-to-noise ratio (SNR) with time. In [8], the authors aimed to more precisely manage the transmission rates in CSMA/CA wireless networks by employing MAB algorithms. The authors suggested using the logarithmic values of the transmission rates of MCS levels rather than their nominal values in the computation of their anticipated throughput measures. The authors introduced two multi-armed bandit (MAB) algorithms embracing logarithmic rates and a sliding window in the course of time. The proposed MAB algorithms were also developed to support frame aggregation schemes known in wireless network standards for elevated throughput, such as IEEE 802.11ax. Another report in [9] discourses QoS-guaranteed AMC for mobile-scalable video multicast. An AMC scheme has been introduced to accommodate both spatial variety and temporal instabilities on multicast channels for providing QoS and enhancing resource utility. An AMC optimization problem that maximizes the resource utility while ensuring each user's QoS has been acquired based on MDPs.

The explained recent works mainly discuss the general wireless communication network; nevertheless, some reports also discuss the AMC mechanism in the narrowband internet-of-things (NB-IoT) domain. As reported in [10], the authors targeted the recurrence factor for NB-IoT, which 3GPP introduces to enhance maximum coupling loss (MCL). The authors addressed downlink 2-dimension link adaptation in NB-IoT. They presented a method for link adaptation to reduce depleted resources and operational time for packet transmission by determining the slightest achievable repetition number and the least potential MCS level. The generative adversarial network (GAN) models are introduced in [11] to manage the stochastic time-stamps of traffic scheduling associated with adaptive MCS values, repetitions, and number of PRBs. One of the limitations of the GAN models is that the system needs to do the periodic re-training to adapt to the network changes.

Some reports also highlight the effect of implementing LA and AMC to achieve the ultra-reliable low latency communications (URLLCs) requirement. The mentioned works are reported in [12–14]. The reports utilize well-known key performance indicators (KPI), such as BLER, throughput, and latency. Nevertheless, the majority of vehicular networks in the reports utilize a roadside unit (RSU) to govern resource management. In this work, we want to investigate the AMC within the autonomous situation as in the 3GPP mode-4 mechanism.

Regarding vehicular networks, especially in cellular communication, the AMC technique is used to be susceptible to particular environmental information. This information includes channel quality indicator (CQI), network latency, block error rate (BLER), and throughput. The report in [15,16] investigated the AMC method for a massive multiple-input multiple-output (MIMO) scenario. Although not discussing vehicular networks in particular, the reports examine the general idea of adapting the AMC in the MIMO and

5G-MIMO systems, respectively. The use of CQI information to construct the AMC mechanism is presented in some reports, such as in [17–21]. From our interpretation, despite the majority of the reports posing a real-world configuration, most of the following reports still need to present a discussion of the overall packet received ratio (PRR) performance.

The real-time adaptation of the MCS has mainly been suggested to implement distributed congestion control in C-V2X, as indicated by the 3GPP and relevant scholarly sources. Additionally, it has been recommended as a viable remedy for transmitting packets of varying sizes, although to a lesser degree [22]. Nevertheless, related research in this field has primarily examined the effects of various fixed MCS configurations, which do not adhere to the requirements set by the European Telecommunications Standards Institute (ETSI) [23]. Alternatively, other studies have primarily discussed the possibility of MCS adaptation without thoroughly assessing its functionality and effectiveness. The authors briefly examine the use of MCS adaptation for congestion management in C-V2X Mode 4 in references [24,25]. They also assess the possible reduction in congestion by raising the MCS, which leads to a fall in performance but a decrease in necessary resources. Nevertheless, these studies need to examine the potential consequences of incorporating this mechanism into the functioning of C-V2X SB-SPS, a significant factor in its overall performance.

The research in reference [26] comprehensively examines various congestion management methods in the context of C-V2X communication. The study also investigates the influence of different MCS configurations on the performance of C-V2X in vehicle congestion. Nevertheless, it should be noted that the research assumes that the MCS configurations stay constant during the simulations. Additionally, it needs to analyze how the MCS adaptation operates regarding congestion management, specifically following the standards established by the ETSI [23]. These standards require that the MCS adaptation perform dynamically throughout continuous transmissions and in response to the congestion circumstances present in the channel.

Additional scholarly works, including references [27,28], explore the feasibility of using MCS adaptation as a potential resolution for transmitting vehicular packets of varying sizes via C-V2X technology. These studies provide a comprehensive examination of the consequences of MCS adaptation and underscore the significance of this method for the specific context in question. Nevertheless, the current models need to include the functioning of MCS adaptation, which is crucial for evaluating the effectiveness of this method in scenarios where packet sizes vary.

The use of MCS adaptability in C-V2X may extend beyond the scope of existing research on the subject. Several studies show this phenomenon using the wireless 802.11p vehicular standard. As shown in references [29,30], the use of MCS adaptation has proven effective in enhancing the efficiency of broadcasts via the dynamic adjustment of the MCS to accommodate changing channel circumstances. The authors in reference [31] emphasize the significance of MCS adaptability concerning the effectiveness of vehicle unicast transmissions, a characteristic of NR-V2X as specified in 5G Release 16 [2].

Despite the significance of evaluating the impacts of implementing the adaptive MCS in the vehicle cellular sidelink, this matter has yet to be explored in scholarly literature. This research aims to tackle this difficulty by thoroughly examining the impacts of the MCS adaptation on the functioning of the C-V2X sidelink. This study specifically examines the effects of MCS adaptation on the resource allocation of C-V2X, the calculation of power levels in the channel, and their implications for the functioning of SB-SPS. This study conducts a comprehensive analysis and precise modeling of the MCS adaptation's performance implications. To the authors' knowledge, this particular area of research has yet to be previously investigated. Ultimately, this research establishes the fundamental principles for future studies seeking to use MCS adaptability in many contexts.

This work utilizes the overall PRR statistic over time as a reinforcement learning (RL) input. While the statistical PRR will be calculated on each packet transmission, we assume that the following network conditions, such as channel information and the signal-to-interference and noise ratio (SINR), are already calculated and carried out in the PRR

records. The following new MCS level will be decided based on the subsequent PRR records. Since the overall PRR record represents the given network and environmental information, our approach will be able to satisfy the real-world scenario.

In this work, we proposed the application of the RL algorithm, as we used the Q-learning approach to adaptively change the MCS index according to the current environmental states. The main contributions in this work can be described as follows:

- We conducted several simulations and showed that static MCS index values had yet to achieve optimal performance.
- We proposed a design of the Q-Learning RL approach to override the MCS index value dynamically.
- Our RL algorithm approach was built on a MATLAB-based simulator and learned as the simulation ran within the specified time limit (T).
- Performance measurement was carried out by comparing PRR results for each static MCS index value scenario with our proposed RL approach.

The remainder of the paper can be explained as follows: Section 2 provides an explanation of the technology. Section 3 describes the proposed RL approach. Section 4 describes the simulation configuration and simulation results. Section 5 provides the conclusion of this work.

## 2. C-V2X Mode 4

### 2.1. Physical Layer

C-V2X employs single carrier frequency division multiple access (SC-FDMA) as its underlying physical layer. Both time division duplex (TDD) and frequency division duplex (FDD) are supported by this system. The LTE frame structure encompasses many temporal units, including the frame, subframe, slot, and symbol. The duration of a single time frame is ten milliseconds, whereas a subframe has a width of one millisecond. Each time slot inside the subframe occupies 0.5 milliseconds. The smallest LTE radio resource unit given to the user equipment is known as a resource block (RB). Each radio band has a minimum frequency width of 180 kHz. In the temporal dimension, an RB spans a duration of one time slot and accommodates a total of seven symbols. Within the area of frequency analysis, an RB is partitioned into subcarriers measuring  $12 \times 15$  kHz. The detailed illustration of an LTE frame can be seen in Figure 1.

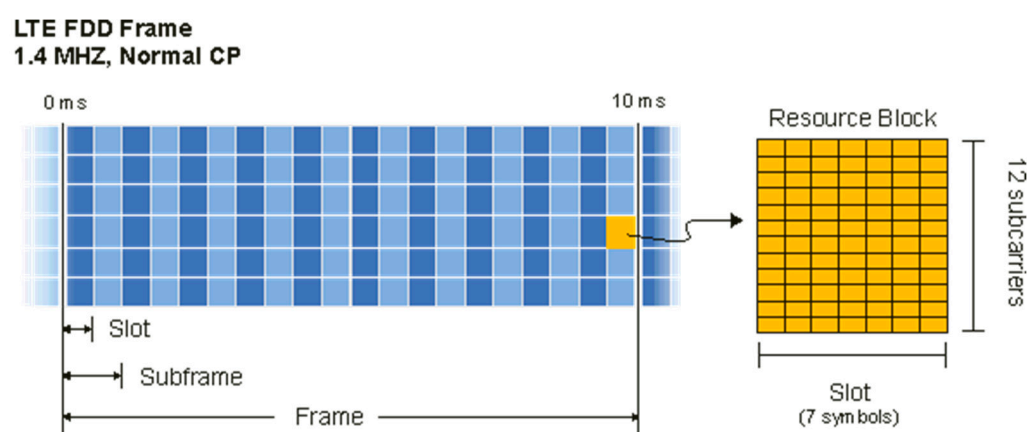


Figure 1. LTE frame [32].

The quantity of resource blocks (RBs) is determined by the channel bandwidth within which the system is operational. The channel bandwidth refers to the officially used channel quantified in megahertz (MHz). It is, therefore, synonymous with the channel spacing. The transmission bandwidth configuration, expressed in RB units, indicates the maximum number of RBs that may be sent over any given channel bandwidth. The transmission bandwidth, sometimes referred to as RBs, is the aggregate number of RBs allocated for a

single transmission. The range of acceptable RBs for a given channel bandwidth might vary from one RB to the maximum allowable RBs [33].

## 2.2. C-V2X Resource Management

The modulation coding scheme (MCS) is associated with the modulation order, which refers to the extent of modulation, such as QPSK, 16QAM, 64QAM, and 256QAM. The modulation order parameter denoted as  $Q_m$  in 3GPP is used to specify the modulation order. The MCS index ( $I_{MCS}$ ) determines the maximum number of useful bits sent per resource element (RE). The performance of C-V2X communication is contingent upon the quality of the radio connection. The higher the MCS index, the greater the RB allocation, increasing the transmission capacity for vital data.

Furthermore, it should be noted that a reciprocal relationship exists between radio conditions and the  $I_{MCS}$  used. Specifically, when radio circumstances deteriorate, the  $I_{MCS}$  is downgraded, decreasing the capacity to transmit vital data. In other terms, the performance of the MCS is contingent upon the frequency of errors occurring. In general, it is customary to establish an error probability threshold that is equivalent to 10%. The  $I_{MCS}$  is modified to ensure the error probability remains under a specified threshold, even when faced with fluctuating radio circumstances. The  $I_{MCS}$  is afterward used to ascertain the transport block size (TBS) estimation. The MCS index, TBS index ( $I_{TBS}$ ), and the associated modulation order ( $Q_m$ ) value have been standardized by the 3GPP [34].

Upon calculating the  $I_{TBS}$  based on the  $I_{MCS}$ , the value of  $N_{PRB}$ , which represents the number of physical resource blocks required to transmit one unit of beacon data (in bits), is determined. The 3GPP provided a reference table that outlines the TBS indices and the corresponding number of physical resource blocks based on the beacon data sizes expressed in bits. The combination of the TBS index, number of physical resource blocks, and beacon data size was cited by the 3GPP. The whole beacon payload in the Mode 4 configuration consists of two components: the beacon data transport block (TB) and the sidelink control information (SCI). These two segments of the TB and SCI need an independent allocation of resources. Therefore, the total number of  $N_{PRB}$  required for a single complete data transmission equals  $N_{PRB}$  multiplied by 2.

The coding rate (CR), spectral efficiency, and lowest signal-to-interference-plus-noise ratio (SINR) are computed for the chosen number of  $N_{PRB}$  in order to evaluate the data transmission process. The data transmission capacity of a certain number of physical resource blocks (PRBs) is contingent upon the modulation order and code rate. Currently, the C-V2X standard accommodates 16-QAM (Quadrature Amplitude Modulation) and QPSK (Quadrature Phase-shift Keying) modulations, which may be used with various code rate configurations. The expression for the effective coding rate (CR) is as follows:

$$CR = \frac{TB}{N_{PRB} \times S \times P \times Q_m}, \quad (1)$$

where  $TB$  is the transport blocks in bits,  $N_{PRB}$  is the number of physical resource blocks,  $S$  is the number of symbols per subcarrier,  $P$  is the number of subcarriers per PRB, and  $Q_m$  is the modulation order [35].

The concept of spectrum efficiency may be traced back to Shannon's theorem, which mathematically defines the maximum amount of information that can be sent across a communication channel,

$$\hat{C} = B_c \log_2(1 + SNR), \quad (2)$$

where  $\hat{C}$  is the capacity in units of bits per second (bit/s),  $B_c$  is the channel bandwidth in Hertz, and SNR is the signal-to-noise ratio.

Shannon's theorem is generally acknowledged as the prevailing constraint on the maximum capacity of information that may be sent across a communication channel. Hence, an increase in signal power or a decrease in interfering signal strength directly corresponds to an amplification in the capacity of the channel to transmit information. The

information-carrying capacity may be considered unlimited in the absence of noise and interference. According to Shannon's capacity formula, it is observed that the impact of a rising interference level (resulting in a lower signal-to-interference-plus-noise ratio (SINR)) on capacity reduction is more significant than what may be anticipated originally. It is worth noting that doubling the interference level does not result in a halving of the capacity, contrary to expectations. This conceptual understanding justifies the implementation of densely populated cells and frequency reuse. The subsequent rise in interference and its mitigated impact on capacity is counterbalanced by the presence of a more significant number of cells and the ability to accommodate a more extensive user base [35].

Presently, the existing systems function with SINR that are in close proximity to the maximum threshold, with just a marginal difference of a few dB. Various modulation and radio schemes exhibit varying degrees of proximity to the limit, prompting the introduction of two parameters to delineate the performance of these distinct schemes. The performance evaluation of modulation schemes may be effectively quantified by using a significant measure derived from the capacity formula. Hence, the parameter denoted as modulation efficiency (alternatively known as channel spectral efficiency) is defined as follows: The radio system is required to compute the least SINR for the chosen  $N_{PRB}$  in order to ensure effective execution of data transmission. The  $I_{MCS}$  and TB sizes have been found to meet the requirements specified in the 3GPP reference table. The minimum SINR or SINR threshold ( $SINR_{min}$ ) might be stated differently depending on the size of the TB. The minimum SINR can be expressed as,

$$SINR_{min} = 10\log_{10}\left(2^{\frac{\eta_c}{a}} - 1\right), \quad (3)$$

where  $\eta_c$  is the channel spectral efficiency,  $a = 0.4$ , is a constant taken from the 3GPP standard [34]. As the value of  $I_{MCS}$  increases, there is a corresponding increase in the  $SINR_{min}$ . This may be attributed to the increasing significance of the beacon (TB) capacity, resulting in a decrease in  $N_{PRB}$  [34]. The increased capacity of the TB may result in a higher signal power required for effective transmission of data. Consequently, the minimum SINR ( $SINR_{min}$ ) exhibits an upward trend.

### 3. Proposed Reinforcement Learning Approach for Adaptive Resource Management

#### 3.1. Impact of MCS Index

This work simulated the C-V2X network and communications using the LTEV2Vsim version 5.4 simulator [36]. The LTEV2Vsim simulator is based on MATLAB R2022a software designed to investigate the C-V2X resource management research. The radio resource parameters are briefly explained in Table 1. We simulated the 1 km highway scenario with one lane in each direction. The vehicle's speed was set at 120 km/h with a 40 km/h standard deviation. Two kinds of vehicle density were simulated, one with the total number of vehicles,  $n_V = 100$  vehicles, and another one with  $n_V = 200$  vehicles. The simulation time ( $T$ ) was set to 60 s, with the position update time resolution ( $t_r$ ) updated every 0.1 s. The channel bandwidth ( $B_c$ ) was set to 10 MHz. We set the transmission power to 23 dBm for this work.

The packet received ratio (PRR) was used to investigate the overall statistical performance. The PRR performance was observed against the vehicle's 10–150 m distance. The overall PRR performance statistic can be obtained by:

$$PRR = \frac{\text{Number of successfully recieved packets}}{\text{Total number of packets}}, \quad (4)$$

PRR values reflect the statistical data of the total number of successfully received packets divided by the total number of packets. The total number of packets includes the retransmission packets. The transmitting vehicle and the designated receiving vehicle are assigned randomly. For the given simulation time ( $T$ ), we recorded the average total number of packets for  $n_V = 100$  vehicles, which was 60,000 packets, and for the  $n_V = 200$  vehicles scenario, which was 120,000 packets.

**Table 1.** Parameter settings.

Parameters	Values
Simulation time ( $T$ )	60 s
Vehicle position update time ( $t_r$ )	0.1 s
Vehicle's speed	120 km/h
Vehicle's speed standard deviation	40 km/h
Road length	1000 m
Number of lanes	1 per directions
Total number of vehicles ( $n_V$ )	100 vehicles, 200 vehicles
Channel bandwidth ( $B_c$ )	10 Mhz
Transmission power ( $P_{Tx}$ )	23 dBm
Beacon size	190 Bytes
Sensing interval	0.1 s
Modulation Coding Scheme (MCS)	0~28
Probability resource keep ( $P_{rk}$ )	0.8
Q-learning predefined epsilon ( $\epsilon$ )	0.1, 0.9
Q-learning discount factor ( $\gamma$ )	0.1
Q-learning learning rate ( $\alpha$ )	0.1, 0.9

The vehicle position update ( $t_r$ ) was carried out at each 0.1 s. Suppose a packet transmission occurs within the 0.1 s space. In that case, the transmission number will be wrapped up for a more straightforward calculation. Following the 3GPP mode-4 mechanism, every vehicle will be assigned a random resource reselection counter. This reselection counter will then decrease by one for every packet transmission. When the counter reaches zero, the vehicle will be given a chance to select a new radio resource for the following transmissions. The probability resource keep ( $P_{rk}$ ) value dictates the chance of selecting the new radio resource. The higher the  $P_{rk}$  value, the greater the chance to select the same radio resource previously assigned and vice versa. Considering the randomness of the simulation works, we have yet to track the performance of individual vehicles.

In the simulation, an RB matrix (RBtable) represented the LTE frame grid. The RBtable consisted of how many RBs were allocated for transmission according to the MCS index. The radio resource management uses a slot-based approach. Other matrices were also created to accommodate the radio calculation and the position update. Every vehicle selected for a packet transmission was paired with the corresponding transmitting and receiving flag. The transmission direction flag was then used to represent the uplink (UL) and downlink (DL). Every vehicle shared the same RBtable to simplify the mode-4 sensing mechanism purposely.

In this study, we focused on how radio resource management works to provide successful C-V2X communication. The radio resource management emphasized how many resource blocks would be provided to each vehicle in the system, which was governed by the MCS index, to reach the highest PRR.

Some related works used particular MCS index values in their investigation. We conducted our initial investigation referencing some works that we have collected so far. Table 2 shows the corresponding related works and the MCS index value used within the report. The 3GPP stated that the MCS index values range from 0 to 28. Thus, the RL algorithm we used was the MCS index value from 0 to 28 in this work. We investigated various Q-learning parameters in this work. For the epsilon value ( $\epsilon$ ), we used  $\epsilon = 0.1$  and  $\epsilon = 0.9$  to represent the "exploitation" and "exploration" actions, respectively, within the epsilon-greedy mechanism [37]. When the epsilon value is close to zero, the system will never explore but tend to exploit previously known knowledge. On the other hand, when the epsilon value is close to one, the system will prefer to use random action rather than to use the information from past knowledge. Similar to the epsilon, we investigated two sets of learning rate ( $\alpha$ ) value scenarios. The learning rate scenarios were set to  $\alpha = 0.1$  and  $\alpha = 0.9$ . The smaller alpha value meant the Q-value would change slower, and vice versa. For the discount factor ( $\gamma$ ) value, since we aimed to achieve the optimal PRR performance, we only

investigated the discount factor value at the low value at  $\gamma = 0.1$ . Under the simulation parameters, as explained in Table 1 and Figures 2 and 3, are the initial simulation results of the corresponding MCS index from the related works for 100 vehicles and 200 vehicles.

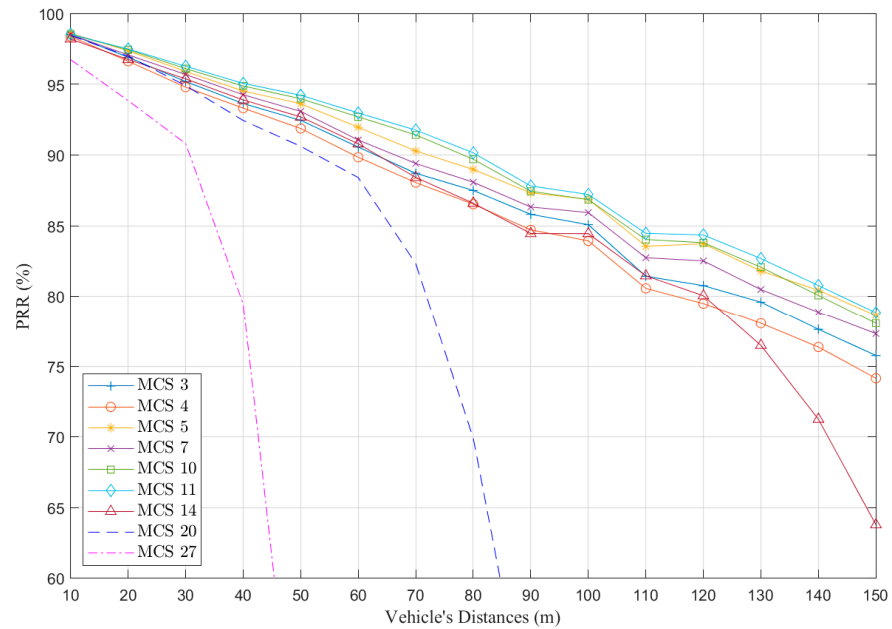


Figure 2. PRR performance of static MCS index value against vehicle's distance with  $n_V = 100$  vehicles.

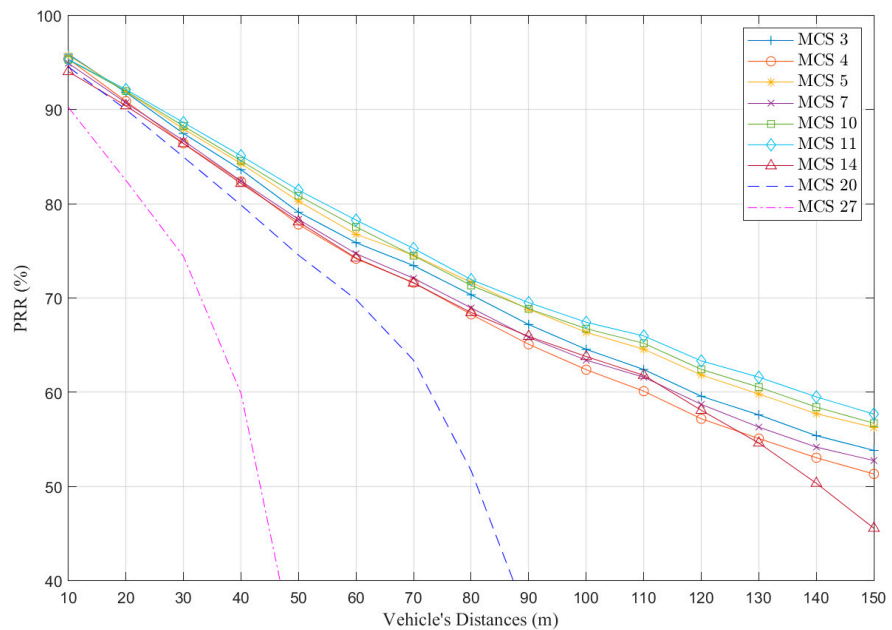


Figure 3. PRR performance of static MCS index value against vehicle's distance with  $n_V = 200$  vehicles.



**Table 2.** MCS index values from various related works.

Related Work	MCS index
[22]	7, 11
[38]	7
[39]	4
[40]	7, 10
[41]	4, 7
[42]	7
[43]	7
[44]	4
[45]	3, 4
[13]	5, 11, 20, 27

Figure 2 shows the PRR performance with corresponding MCS index values based on Table 1 against the vehicle's distance. We conducted our initial simulation as a benchmark foundation to investigate our proposed RL algorithm approach for the adaptive MCS index value later. As shown in Figure 2, the total number of vehicles was set to 100 vehicles. Since the vehicle's density was notably at a fair amount, most of the MCS index values PRR performance were relatively high. For instance, the most significant performance difference was developed on the higher MCS index, 20 and 27. As explained in the previous section, a higher MCS index will result in a larger number of physical resource blocks. The larger number of physical resource blocks required more transmission power. Hence, the higher MCS index suffered poor PRR performance. A similar result was also obtained in Figure 3, where the number of vehicles was higher, at 200 vehicles in total. Identically with Figure 2, the higher MCS index also suffered poor PRR performance. The condition worsened with the higher vehicle number. More vehicles deployed in the system meant more RBs needed to be occupied for data transmission. Thus, when the number of physical RBs is large and the number of vehicles is also high, this may cause a lack of radio resources.

From Figures 2 and 3, we can see that using a high MCS index value will produce low PRR performance. Instead, the overall PRR performance may drop significantly along with the increasing vehicle distance. This situation may be caused by the lack of the transmission power to satisfy the SINR threshold on each TB.

### 3.2. RL Q-Learning Approach

Q-learning, as a RL method, is used to provide the MCS index adaptability following the Q-learning procedure. Figure 4 shows the general RL architecture used in this work. We proposed the continuous "on-the-run" learning process. Thus, the learning time steps were bound to the simulation time (T) setting. The Q-learning obtains the Q-value from the corresponding action and state pair. The action function used in our proposed method was the possible MCS index value within the C-V2X environment. It can be represented as  $A_k$  ( $k = 1, \dots, K$ ), where  $K = 28$ . We designed the state function as the calculated PRR on each time step. The system can select one action on each time step and then give feedback on the MCS index value for the resource management calculation. The resource management allocation will then produce the current time step PRR. The Q-learning function itself can be expressed as:

$$Q(S, A)_t \leftarrow Q(S, A)_{t-1} + \alpha \times \Delta Q_t, \quad (5)$$

where  $t$  is the time steps,  $S$  is the state,  $A$  is the action,  $\alpha \in [0, 1]$  is the learning rate, and  $\Delta Q_t$  is the difference between the value function target and the current Q-value. The value function target and  $\Delta Q_t$  can be expressed as,

$$Y_t = R_t + \gamma \max_A Q(S, A)_t, \quad (6)$$

and

$$\Delta Q_t = Y_t - Q(S, A)_{t-1}, \quad (7)$$

respectively, where  $R_t$  is the reward function in current time step and  $\gamma \in [0 1]$  is the discount factor.

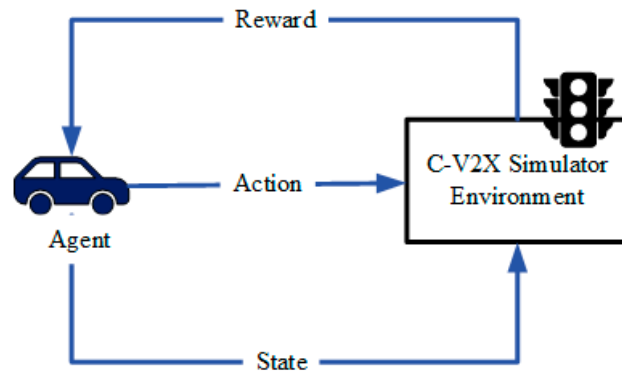


Figure 4. PRR performance for static MCS index value against vehicle’s distance.

The corresponding reward function can be attained as:

$$R_t = PRR_t - PRR_{t-1}, \tag{8}$$

in the initial time step, the initial Q-value is set to equal to current  $PRR_0$  based on the initial MCS index value  $I_{MCS_0}$ . As the learning step runs, the appropriate action was selected according to the  $\epsilon$ -greedy exploration mechanism,

$$\pi_t = \begin{cases} \text{choose random action in } A, & \text{if } \zeta_t < \epsilon \\ \arg \max_A Q(S, A)_t, & \text{otherwise} \end{cases}, \tag{9}$$

where  $\zeta_t$  is a random value between 0 and 1, and  $\epsilon \in [0 1]$  is the predefined epsilon value. Algorithm 1 explains the detailed procedure of our proposed RL approach.

---

**Algorithm 1.** Pseudocode for proposed RL approach

---

- 1: **Initialize:**
  - 2:     initialize  $T, I_{MCS_0}, \alpha, \gamma, \epsilon$
  - 3:     calculate  $PRR_0$  based on  $I_{MCS_0}$
  - 4: **while**  $t \leq T$  **do**
  - 5:   for each: LTE event
  - 6:     calculate the current Q-value based  $PRR_0$
  - 7:     calculate the current rewards based on Equation (8)
  - 8:     explore the next action that satisfy the condition in Equation (9)
  - 9:     calculate the current value function target based on Equation (6)
  - 10:    calculate the current  $\Delta Q_t$  based on Equation (7)
  - 11:    update the Q-value based on Equation (5)
  - 12:    calculate the new  $PRR_t$
  - 13: **end while**
- 

#### 4. Simulation Results

The following section discusses the simulation results for the static MCS index as a benchmark compared to the proposed RL approach for the adaptive MCS index. As mentioned before, a higher MCS index will result in higher transport block capacity. However, the larger capacity means higher power consumption. On the other hand, using a lower MCS index does not mean the overall system performance will become the finest. A lower MCS index will result in a smaller transport block capacity. Although it may cause a lower power consumption, a small transport block may result in higher resource blocks allocation. The total number of resource blocks is bound to the corresponding channel

bandwidth, as stated by the 3GPP. Hence, if the number of resource blocks allocated to each vehicle is massive, then it will lead to a need for more radio resources.

Figure 5 shows that the MCS index 11 achieved the best PRR performance among the other static MCS index values. This figure shows that the MCS index 11 might be sufficient for the corresponding environment. The transmission power amount was sufficient overall to provide the SINR threshold for the given MCS index 11. However, more was needed to satisfy the 90% PRR performance. The MCS index 11 PRR performance plunged below 90% at the vehicle’s distance of 80 m and beyond. In order to satisfy the demand for reliability and a low latency trend, the acceptable PRR is generally above 90%. Figure 5 also shows that our proposed RL approach maintained a 90% PRR performance up to 130 m of the vehicle’s distance. The total number of vehicles in the first scenario was set to 100 vehicles. Our proposed RL approach has proven capable of adaptively learning the vehicle’s environmental situations and adjusting the corresponding MCS index. MCS index adjustment is carried out by comparing the actual PRR with the past steps PRR. Thus, the “on-the-run” methods may continuously learn. Our proposed method’s overall PRR performance improvement was about 10% for the  $n_V = 100$  vehicles scenario.

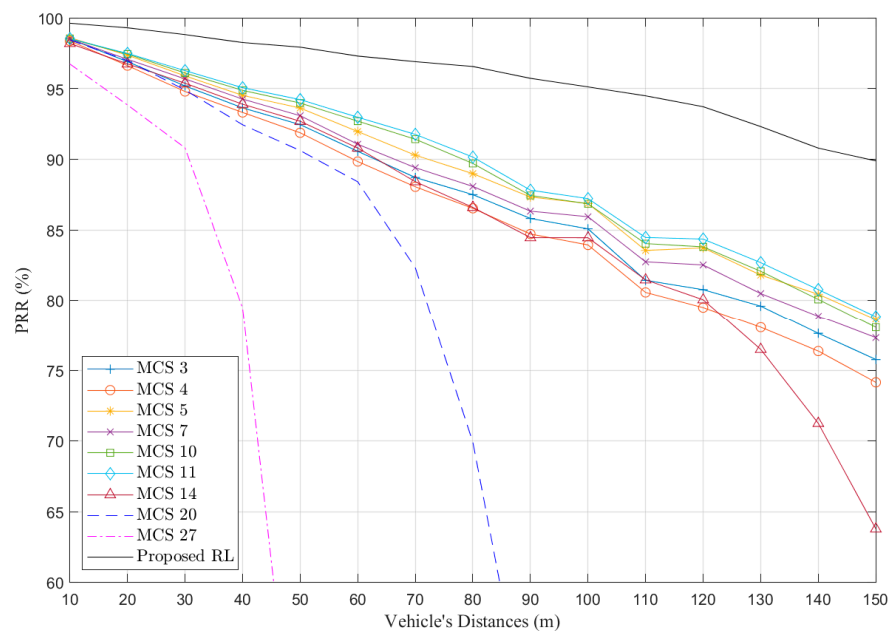


Figure 5. Overall PRR performance comparison for  $n_V = 100$  vehicles scenario.

The overall PRR performance comparison for the  $n_V = 200$  vehicles scenario is shown in Figure 6. From Figure 6, it can be seen that the MCS index 11 achieved the best overall performance from the other static MCS index values. Nevertheless, the static MCS indexes were not performing well under the massive vehicle number. Even the MCS index 11 plunged under 90% PRR performance in just around 25 m of vehicle distance. Despite the overall performance in the “high-density” scenario still needing improvements, our proposed RL approach achieved much better than the static MCS index. The RL approach could satisfy the 90% PRR performance up to 60 m of the vehicle’s distance under a high-density scenario. From both Figures 5 and 6, we can observe that the performance from the high MCS indexes, 20 and 27, were deficient. The case of MCS 20 and 27 may be caused by a lack of transmission power in order to satisfy the size of the large chunk of transport blocks carried by the corresponding MCS index. The large transport blocks size may have been of benefit in the large beacon size scenario. Focusing on the very beginning of Figure 4, the MCS index 20 achieved the highest PRR performance. However, the performance immediately plunged significantly as the vehicle’s distance increased. For the  $n_V = 200$  vehicles scenario, our proposed methods achieved an overall PRR performance improvement of 21%.

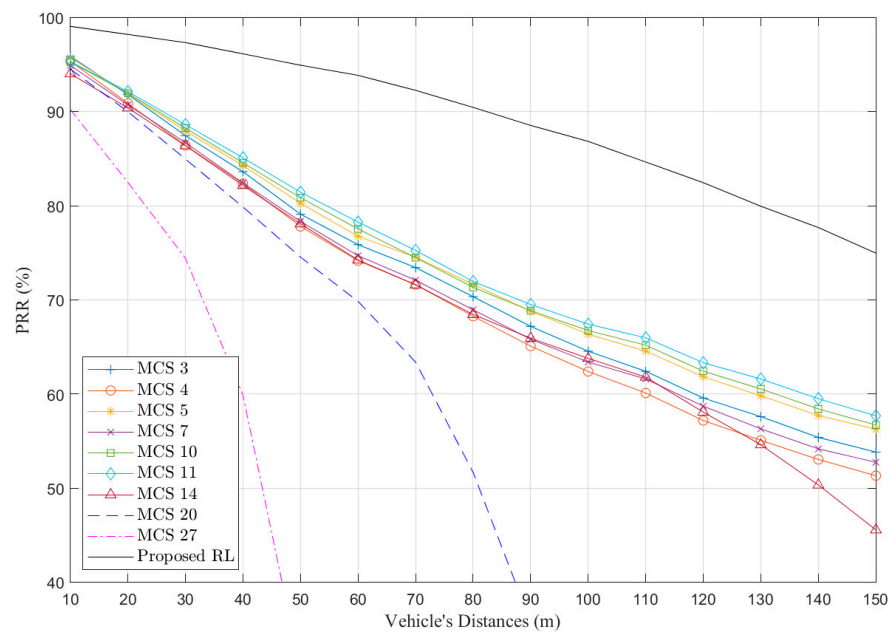


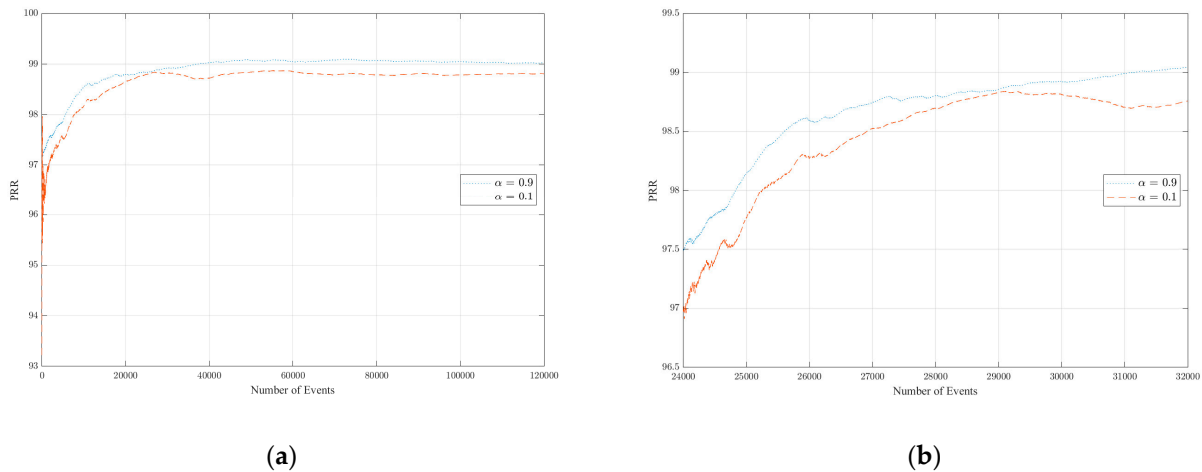
Figure 6. Overall PRR performance comparison for  $n_V = 200$  vehicles scenario.

In order to investigate the proposed RL approach, Figures 7 and 8 show the PRR and rewards evolution against the LTE events number. Figure 7 shows the PRR results taken from the different learning rates,  $\alpha = 0.1$  and  $\alpha = 0.9$ . The lower alpha means the system will prefer not to refrain from learning anything from the newly selected action (MCS index). The higher alpha value will make the system value the more recent information, thus resulting in a faster Q-value change. It can be seen from Figure 7b that the higher alpha value achieved better results utilizing the new information. Although the PRR difference was slim, the higher alpha value showed more promising results. Figure 7a shows that the proposed RL methods achieved a stable performance around 40,000 events, each event representing one packet transmission. Assuming a uniformly random distribution packet generation, the convergence time shown in Figure 7a was around 20 s. The estimated convergence time achieved by

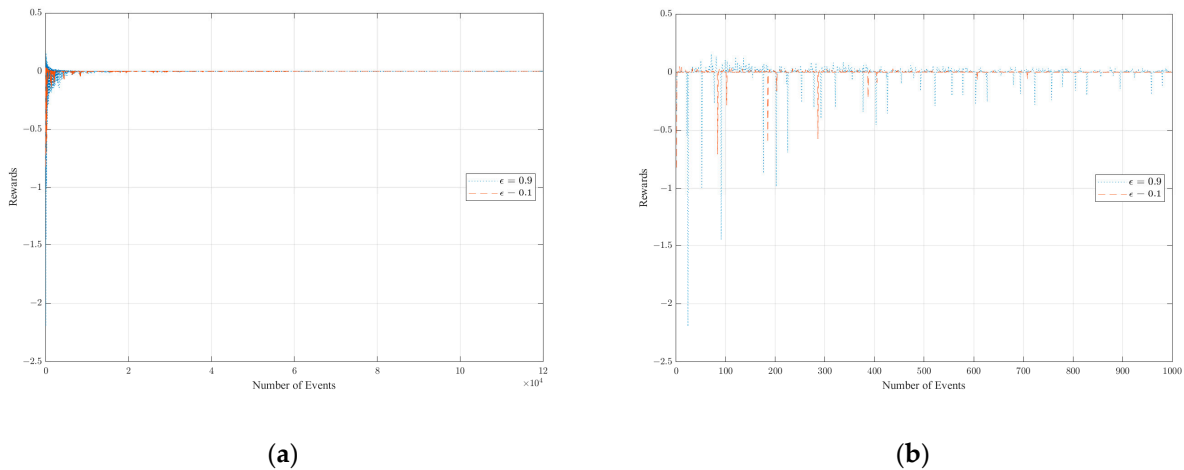
$$T_{converge} = \frac{\text{converged event number}}{\text{total number of transmitted packets}} \times T, \tag{10}$$

where  $T = 60$  s is the simulation time, total number of transmitted packets = 120,000 packets, and the converged event number = 40,000.

Figure 8 shows the rewards evolution against the number of events. Similar to Figure 7, the rewards evolution results were taken from two different epsilon values,  $\epsilon = 0.1$  and  $\epsilon = 0.9$ . As we can see from Figure 8b, the results of  $\epsilon = 0.9$  show more scattered points compared to the results of  $\epsilon = 0.1$ . The lower epsilon will result in the RL not exploring the possible new actions; in other words, the RL will tend to exploit the already-known knowledge. The exploring behavior can be seen in the more scattered points of the results of  $\epsilon = 0.9$ . As the RL explored every possible action provided, it resulted in various performance states. These various states caused more rewards to be produced in the learning episode, resulting in more scattered points.



**Figure 7.** PRR evolution against number of events from the proposed RL approach; (a) The overall PRR evolution from  $n_V = 200$  vehicles scenario; (b) The detailed PRR evolution for the first 24,000 to 32,000 events.



**Figure 8.** Rewards evolution against number of events from the proposed RL approach; (a) The overall rewards evolution from the  $n_V = 200$  vehicles scenario; (b) The detailed rewards evolution for the first 1000 events.

### 5. Conclusions

The impacts of the variable and adaptive MCS index in the C-V2X communication should be discussed among researchers. In this work, we propose the application of the RL algorithm, as we used the Q-learning approach to adaptively change the MCS index according to the current environmental states. We gathered some MCS index values used by related works as benchmark information, and we then used the MCS index values in our initial simulations. Our proposed RL approach design was implemented as an “on-the-run” learning process. The simulation results showed that our proposed RL approach outperformed the static MCS index for the total number of vehicles, being 100 vehicles and 200 vehicles. Our proposed method achieved an average of 10% and 21% PRR improvements for the  $n_V = 100$  and  $n_V = 200$  vehicles scenario, respectively. Our proposed method also showed that the learning mechanism reaches its stability at an acceptable number of events. Despite the promising results, this work only utilizes the PRR statistics, and it is possible to improve in the case of real-world parameter configuration. These real-world parameters will be carried out in our future works.

**Author Contributions:** Conceptualization, Y.-F.H. and J.-K.C.; Investigation, T.I.B. and Y.-F.H.; Methodology, T.I.B. and Y.-F.H.; Software, T.I.B. and J.-K.C.; Supervision, Y.-F.H. and J.-K.C.; Writing—Original Draft, T.I.B. and Y.-F.H.; Writing—Review and Editing, T.I.B. and Y.-F.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the National Science and Technology Council of Taiwan under grant NSTC-111-2221-E-324-018.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Häfner, B.; Bajpai, V.; Ott, J.; Schmitt, G.A. A Survey on Cooperative Architectures and Maneuvers for Connected and Automated Vehicles. *IEEE Commun. Surv. Tutor.* **2022**, *24*, 380–403. [\[CrossRef\]](#)
- Varatharaajan, S.; Grossmann, M.; Galdo, G.D. 5G New Radio Physical Downlink Control Channel Reliability Enhancements for Multiple Transmission-Reception-Point Communications. *IEEE Access* **2022**, *10*, 97394–97407. [\[CrossRef\]](#)
- Sehla, K.; Nguyen, T.M.T.; Pujolle, G.; Velloso, P.B. Resource Allocation Modes in C-V2X: From LTE-V2X to 5G-V2X. *IEEE Internet Things J.* **2022**, *9*, 8291–8314. [\[CrossRef\]](#)
- Roux, P.; Sesia, S.; Mannoni, V.; Perraud, E. System level analysis for ITS-G5 and LTE-V2X performance comparison. In Proceedings of the 2019 IEEE 16th International Conference on Mobile Ad Hoc and Smart Systems, MASS 2019, Monterey, CA, USA, 4–7 November 2019; pp. 1–9.
- Ansari, S.; Alnajjar, K.A. Adaptive Modulation and Coding: A Brief Review of the Literature. *Int. J. Commun. Antenna Propag.* **2023**, *13*, 43–54. [\[CrossRef\]](#)
- Elwekeil, M.; Wang, T.; Zhang, S. Deep learning based adaptive modulation and coding for uplink multi-user SIMO transmissions in IEEE 802.11ax WLANs. *Wirel. Netw.* **2021**, *27*, 5217–5227. [\[CrossRef\]](#)
- Oh, J.E.; Jo, A.M.; Jeong, E.R. MCS Selection Based on Convolutional Neural Network in TDD System. *Int. J. Electr. Electron. Res.* **2023**, *11*, 485–489. [\[CrossRef\]](#)
- Cho, S. Use of Logarithmic Rates in Multi-Armed Bandit-Based Transmission Rate Control Embracing Frame Aggregations in Wireless Networks. *Appl. Sci.* **2023**, *13*, 8485. [\[CrossRef\]](#)
- Jiang, Q.; Leung, V.C.M.; Tang, H. QoS-Guaranteed Adaptive Modulation and Coding for Wireless Scalable Video Multicast. *IEEE Trans Circuits Syst Video Technol* **2022**, *32*, 1696–1700. [\[CrossRef\]](#)
- Sanei, F.; Farbeh, H. A link adaptation scheme for reliable downlink communications in narrowband IoT. *Microelectron. J.* **2021**, *114*, 105154. [\[CrossRef\]](#)
- Karmakar, R.; Kaddoum, G.; Chattopadhyay, S. SmartCon: Deep Probabilistic Learning-Based Intelligent Link-Configuration in Narrowband-IoT Toward 5G and B5G. *IEEE Trans. Cogn. Commun. Netw.* **2022**, *8*, 1147–1158. [\[CrossRef\]](#)
- Huang, Y.; Thomas Hou, Y.; Lou, W. DELUXE: A DL-Based Link Adaptation for URLLC/eMBB Multiplexing in 5G NR. *IEEE J. Sel. Areas Commun.* **2022**, *40*, 143–162. [\[CrossRef\]](#)
- Gao, Y.; Yang, H.; Hong, X.; Chen, L. A Hybrid Scheme of MCS Selection and Spectrum Allocation for URLLC Traffic under Delay and Reliability Constraints. *Entropy* **2022**, *24*, 727. [\[CrossRef\]](#)
- Nayak, S.; Roy, S. Novel Markov Chain Based URLLC Link Adaptation Method for 5G Vehicular Networking. *IEEE Trans. Veh. Technol.* **2021**, *70*, 12302–12311. [\[CrossRef\]](#)
- Liao, Y.; Yang, Z.; Yin, Z.; Shen, X. DQN-Based Adaptive MCS and SDM for 5G Massive MIMO-OFDM Downlink. *IEEE Commun. Lett.* **2023**, *27*, 185–189. [\[CrossRef\]](#)
- Radbord, A.; Harsini, J.S. Slow and fast adaptive modulation and coding for uplink massive MIMO systems with packet retransmission. *IET Commun.* **2022**, *16*, 915–928. [\[CrossRef\]](#)
- Parsa, A.; Moghim, N.; Salavati, P. Joint power allocation and MCS selection for energy-efficient link adaptation: A deep reinforcement learning approach. *Comput. Netw.* **2022**, *218*, 109386. [\[CrossRef\]](#)
- Geiser, F.; Wessel, D.; Hummert, M.; Weber, A.; Wübben, D.; Dekorsy, A.; Viseras, A. DRLLA: Deep Reinforcement Learning for Link Adaptation. *Telecom* **2022**, *3*, 692–705. [\[CrossRef\]](#)
- Khan, J.; Jacob, L. Cognitive Sub-Band Scheduling and Link Adaptation for 5G URLLC. *IEEE Trans. Cogn. Commun. Netw.* **2021**, *7*, 1280–1290. [\[CrossRef\]](#)
- Han, N.; Kim, I.M.; So, J. Lightweight LSTM-Based Adaptive CQI Feedback Scheme for IoT Devices. *Sensors* **2023**, *23*, 4929. [\[CrossRef\]](#)
- Ye, X.; Yu, Y.; Fu, L. Deep Reinforcement Learning Based Link Adaptation Technique for LTE/NR Systems. *IEEE Trans. Veh. Technol.* **2023**, *72*, 7364–7379. [\[CrossRef\]](#)
- Burbano-Abril, A.; McCarthy, B.; Lopez-Guerrero, M.; Rangel, V.; O'Driscoll, A. MCS Adaptation within the Cellular V2X Sidelink. In Proceedings of the 2021 IEEE Conference on Standards for Communications and Networking, CSCN 2021, Virtual, 15–17 December 2021; pp. 111–117.

23. TS 103 574; Intelligent Transport Systems (ITS); Congestion Control Mechanisms for the C-V2X PC5 Interface; Access Layer Part. ETSI, Sophia Antipolis, Cedex: Valbonne, France, 2018. Available online: [https://www.etsi.org/deliver/etsi\\_ts/103500\\_103599/103574/01.01.01\\_60/ts\\_103574v010101p.pdf](https://www.etsi.org/deliver/etsi_ts/103500_103599/103574/01.01.01_60/ts_103574v010101p.pdf) (accessed on 13 October 2023).
24. 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.* **2017**, *12*, 30–39. [[CrossRef](#)]
25. Mansouri, A.; Martinez, V.; Härrä, J. A First Investigation of Congestion Control for LTE-V2X Mode 4. In Proceedings of the 2019 15th Annual Conference on Wireless On-demand Network Systems and Services (WONS), Wengen, Switzerland, 22–24 January 2019; pp. 56–63.
26. Bazzi, A. Congestion Control Mechanisms in IEEE 802.11p and Sidelink C-V2X. In Proceedings of the 2019 53rd Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, 3–6 November 2019; pp. 1125–1130.
27. Bazzi, A.; Zanella, A.; Masini, B.M. Optimizing the Resource Allocation of Periodic Messages with Different Sizes in LTE-V2V. *IEEE Access* **2019**, *7*, 43820–43830. [[CrossRef](#)]
28. Molina-Masegosa, R.; Gozalvez, J.; Sepulcre, M. Comparison of IEEE 802.11p and LTE-V2X: An Evaluation with Periodic and Aperiodic Messages of Constant and Variable Size. *IEEE Access* **2020**, *8*, 121526–121548. [[CrossRef](#)]
29. Yao, Y.; Zhou, X.; Zhang, K. Density-Aware Rate Adaptation for Vehicle Safety Communications in the Highway Environment. *IEEE Commun. Lett.* **2014**, *18*, 1167–1170. [[CrossRef](#)]
30. Haque, K.F.; Abdelgawad, A.; Yanambaka, V.P.; Yelamarthi, K. Lora architecture for v2x communication: An experimental evaluation with vehicles on the move. *Sensors* **2020**, *20*, 6876. [[CrossRef](#)]
31. Camp, J.; Knightly, E. Modulation Rate Adaptation in Urban and Vehicular Environments: Cross-Layer Implementation and Experimental Evaluation. *IEEE/ACM Trans. Netw.* **2010**, *18*, 1949–1962. [[CrossRef](#)]
32. LTE Physical Layer Overview. Available online: [https://rfmw.em.keysight.com/wireless/helpfiles/89600b/webhelp/subsystems/lte/content/lte\\_overview.htm](https://rfmw.em.keysight.com/wireless/helpfiles/89600b/webhelp/subsystems/lte/content/lte_overview.htm) (accessed on 28 August 2023).
33. Rumney, M. *LTE and the Evolution to 4G Wireless: Design and Measurement Challenges*, 2nd ed.; John Wiley & Sons Limited: West Sussex, UK, 2013.
34. TS 36.213 V14.3.0; 3rd Generation Partnership Project; Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Layer Procedures (Release 14). 3GPP, Sophia Antipolis: Valbonne, France, 2017.
35. Steer, M. *Microwave and RF Design Radio Systems*; NC State University: Raleigh, NC, USA, 2019.
36. Cecchini, G.; Bazzi, A.; Masini, B.M.; Zanella, A. LTEV2Vsim: An LTE-V2V Simulator for the Investigation of Resource Allocation for Cooperative Awareness. In Proceedings of the 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems, Napoli, Italy, 26–28 June 2017; pp. 80–85.
37. Zhou, X. Optimal Values Selection of Q-learning Parameters in Stochastic Mazes. *J. Phys. Conf. Ser.* **2022**, *2386*, 012037. [[CrossRef](#)]
38. Abanto-Leon, L.F.; Koppelaar, A.; de Groot, S.H. Enhanced C-V2X Mode-4 Subchannel Selection. In Proceedings of the 2018 IEEE 88th Vehicular Technology Conference (VTC-Fall), Chicago, IL, USA, 27–30 August 2018; pp. 1–5.
39. Kang, B.; Jung, S.; Bahk, S. Sensing-Based Power Adaptation for Cellular V2X Mode 4. In Proceedings of the 2018 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN), Seoul, Republic of Korea, 22–25 October 2018; pp. 1–4.
40. Campolo, C.; Molinaro, A.; Romeo, F.; Bazzi, A.; Berthet, A.O. Full duplex-aided sensing and scheduling in cellular-V2X mode 4. In Proceedings of the International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc), Catania, Italy, 2–5 July 2019; pp. 19–24.
41. Haider, A.; Hwang, S.H. Adaptive Transmit Power Control Algorithm for Sensing-Based Semi-Persistent Scheduling in C-V2X Mode 4 Communication. *Electronics* **2019**, *8*, 846. [[CrossRef](#)]
42. Yoon, Y.; Kim, H. Resolving persistent packet collisions through broadcast feedback in cellular V2X communication. *Future Internet* **2021**, *13*, 211. [[CrossRef](#)]
43. Yang, J.M.; Yoon, H.; Hwang, S.; Bahk, S. PRESS: Predictive Assessment of Resource Usage for C-V2V Mode 4. In Proceedings of the 2021 IEEE Wireless Communications and Networking Conference (WCNC), Nanjing, China, 29 March–1 April 2021; pp. 1–6.
44. Hirai, T.; Wakamiya, N.; Murase, T. NOMA-dependent Low-Powered Retransmission in Sensing-based SPS for Cellular-V2X Mode 4. In Proceedings of the IEEE Vehicular Technology Conference, London, UK, 26–29 September 2022.
45. Yin, J.C.; Hwang, S.H. Variable MCS method for LTE V2V Mode4. In Proceedings of the 2021 International Conference on Information and Communication Technology Convergence (ICTC), Jeju Island, Republic of Korea, 20–22 October 2021; pp. 1368–1370.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.