

Article

Joint Optimization of Age of Information and Energy Consumption in NR-V2X System Based on Deep Reinforcement Learning

Shulin Song ^{1,2,†}, Zheng Zhang ^{1,2,†}, Qiong Wu ^{1,2,*} , Pingyi Fan ³  and Qiang Fan ⁴ 

¹ School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China; songshulin@jiangnan.edu.cn (S.S.); zhengzhang@stu.jiangnan.edu.cn (Z.Z.)

² Zhuhai Fudan Innovation Institute, Zhuhai 519031, China

³ Department of Electronic Engineering, Beijing National Research Center for Information Science and Technology, Tsinghua University, Beijing 100084, China; fpy@tsinghua.edu.cn

⁴ Qualcomm, San Jose, CA 95110, USA; qf9898@gmail.com

* Correspondence: qiongwu@jiangnan.edu.cn; Tel.: +86-0510-8591-0633

† These authors contributed equally to this work.

Abstract: As autonomous driving may be the most important application scenario of the next generation, the development of wireless access technologies enabling reliable and low-latency vehicle communication becomes crucial. To address this, 3GPP has developed Vehicle-to-Everything (V2X) specifications based on 5G New Radio (NR) technology, where Mode 2 Side-Link (SL) communication resembles Mode 4 in LTE-V2X, allowing direct communication between vehicles. This supplements SL communication in LTE-V2X and represents the latest advancements in cellular V2X (C-V2X) with the improved performance of NR-V2X. However, in NR-V2X Mode 2, resource collisions still occur and thus degrade the age of information (AOI). Therefore, an interference cancellation method is employed to mitigate this impact by combining NR-V2X with Non-Orthogonal multiple access (NOMA) technology. In NR-V2X, when vehicles select smaller resource reservation intervals (RRIs), higher-frequency transmissions use more energy to reduce AoI. Hence, it is important to jointly consider AoI and communication energy consumption based on NR-V2X communication. Then, we formulate such an optimization problem and employ the Deep Reinforcement Learning (DRL) algorithm to compute the optimal transmission RRI and transmission power for each transmitting vehicle to reduce the energy consumption of each transmitting vehicle and the AoI of each receiving vehicle. Extensive simulations demonstrate the performance of our proposed algorithm.

Keywords: new radio vehicle-to-everything; deep reinforcement learning; non-orthogonal multiple access; age of information; energy consumption



Citation: Song, S.; Zhang, Z.; Wu, Q.; Fan, P.; Fan, Q. Joint Optimization of Age of Information and Energy Consumption in NR-V2X System Based on Deep Reinforcement Learning. *Sensors* **2024**, *24*, 4338. <https://doi.org/10.3390/s24134338>

Academic Editor: Jordi Mongay Batalla

Received: 8 June 2024

Revised: 28 June 2024

Accepted: 2 July 2024

Published: 4 July 2024



Copyright: © 2024 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

As autonomous driving is one of the most promising application fields for the next generation of communication systems, the development of reliable and low-latency vehicle communication becomes crucial. Such technologies not only enhance interconnectivity between vehicles but also facilitate efficient communication between vehicles and infrastructure. For autonomous driving vehicles, wireless access technologies play a critical role by providing real-time information exchange and collaboration capabilities among vehicles, thereby enhancing driving safety and efficiency. Therefore, continuous innovation and development in wireless access technologies hold strategic significance, further propelling the advancement and adoption of autonomous driving vehicle technology [1,2]. In recent related research, the scheduling and allocation of factors that affect communication performance are the main research directions [3–7]. The main research method focuses on DRL [8–10].

3GPP has formulated V2X specifications based on 5G NR technology to support ultra-low latency and ultra-high reliability in evolving vehicle applications, communications, and service requirements [11]. As pointed out in [12], the development of SL in NR-V2X is to supplement and expand the SL communication in LTE-V2X. However, the autonomous resource allocation method used in Mode 2 still suffers from resource collisions. When the RRI decreases and the vehicle occupies more resources, the collision probability gradually increases. Collisions mean one vehicle receiving multiple messages simultaneously, which results in the mutual interference between these messages and directly reduces the Signal-to-Interference-plus-Noise Ratio (SINR) of each message, which degrades the transmission time. Therefore, NOMA is used to mitigate the impact of this situation. When the vehicle receives multiple collision messages, NOMA decodes these messages separately to increase the SINR of messages with relatively low power and improve communication performance.

Furthermore, as mentioned in [13], while ensuring communication effectiveness, energy consumption also needs to be considered. The transmission power also affects the SINR and energy consumption during the transmission process [14]. When the power is high, the SINR is more likely to meet the requirements for successful transmission, thereby reducing AoI, but energy consumption will also increase at the same time [15]. Therefore, in NR-V2X communication, it is necessary to comprehensively consider the balance between communication effectiveness and energy consumption. In order to address this issue, a resource allocation scheme based on DRL is proposed to allocate RRI and power for vehicles ensuring low energy consumption and low information age of the system during the communication process (The source code has been released at the following link: <https://github.com/qiongwu86/Joint-Optimization-of-AoI-and-Energy-Consumption-in-NR-V2X-System-based-on-DRL> accessed on 6 June 2024). The performance of our proposed resource allocation method is evaluated through simulation experiments, and the results demonstrate that it can improve the communication performance of the NR-V2X vehicle networking system.

The remainder of this paper is structured as follows: Section 2 provides a review of the related literature. In Section 3, we introduce the system model and formulate the optimization problem. Section 4 simplifies the formulated optimization problem and presents a near-optimal solution using DRL. We conduct simulations to demonstrate the effectiveness of our proposed method in Section 5, followed by concluding remarks in Section 6.

2. Related Work

In this section, we have reviewed the existing work on the analysis, network optimization, and the application of reinforcement learning in NR-V2X systems. Rehman et al. proposed an analytical model for evaluating the NR-V2X communication performance, focusing on the sensor-based semi-persistent scheduling operations defined in NR-V2X Mode 2 and comparing them with LTE-V2X Mode 4. For different physical layer specifications, the average packet success probability for LTE-V2X and NR-V2X was analyzed, and a moment matching approximation method was used to approximate the SINR statistics under the Nakagami-lognormal composite channel model. It was shown that, under conditions of relatively large inter-vehicle spacing and a high number of vehicles, NR-V2X outperforms LTE-V2X [16]. Anwar et al. evaluated and compared the PHY layer performance of various V2V communication technologies. The results showed that NR-V2V outperforms other standards (such as IEEE 802.11bd) in terms of reliability, range, delay, and data rate. However, under the same modulation and coding scheme, IEEE 802.11bd performs better in terms of PER. Although the lowest MCS of NR-V2X is more reliable than IEEE 802.11bd, IEEE 802.11bd has a wider range. Overall, NR-V2X performs best in V2V communication [17].

Ref. [18] investigated the energy consumption and AoI of a device that is a single-source node, considering potential transmission failures due to poor channel conditions from the source node to the receiver. For a threshold-based retransmission strategy in

the system, the corresponding closed-form expressions for average AoI and energy consumption were derived, which can be used to estimate channel failure probabilities and maximum retransmission attempts. Authors of [19] adopted the Truncated Automatic Repeat Request scheme, where terminal devices repeatedly send the current status update until reaching the maximum allowable transmission attempts or generating a new status update. Closed-form expressions for average AoI, average peak AoI, and average energy consumption are derived based on the evolution process of AoI. Authors of [20] primarily considered scenarios where multiple information sources are needed to transmit information for completing status updates, thereby reducing AoI. It investigated the problem of packet scheduling based on information freshness for application-oriented scenarios with correlated multiple information sources. Specifically, it employs AoI to characterize the freshness of status updates for applications and formulates the application-oriented scheduling problem as an MDP problem, utilizing DRL for solving.

Liang et al. proposed an implementation method for an Integrated Sensing and Communication (ISAC) system for vehicular networks, addressing the potential performance degradation caused by the coexistence of millimeter-wave radar and communication by extending NR-V2X Mode 2. By using semi-persistent scheduling (SPS) resource selection and dynamically adjusting the radar scanning cycle and transmission power of each vehicle based on the speed and channel congestion status reported by neighboring vehicles, the ISAC system ensures that high-priority vehicles occupy spectrum resources preferentially. Simulation results validated the effectiveness of this approach in improving radar and communication performance, and the ISAC system could better coordinate the coexistence of radar and communication functions, improving the overall performance and security of vehicular networks [21]. Song et al. proposed a scheme for SL resource allocation in NR-V2X based on 5G cellular mobile communication networks in Mode 1. By using hybrid spectrum access technology and periodic reporting of channel state information, SL resource allocation was modeled as a mixed binary integer nonlinear programming problem to maximize the total throughput of NR-V2X networks among different subcarriers while complying with total available power and minimum transmission rate constraints. Simulation results showed that the proposed power allocation scheme could save energy, and the suboptimal SL resource allocation algorithm outperformed other methods [22]. MolinaGalan et al. conducted an in-depth analysis and evaluated the performance of 5G NR-V2X Mode 2 under different traffic patterns. The study pointed out that additional reselections could make SPS more unstable and prone to collisions. Moreover, frequent resource reselections could increase implementation costs. Therefore, they proposed suggestions for an adjusted reevaluation mechanism to reduce implementation costs and improve system performance. This work provided valuable insights for the further development and optimization of NR-V2X Mode 2 [23]. Soleymani et al. focused on the joint energy efficiency and total rate maximization problem of autonomous resource selection in NR-V2X vehicle communication to meet reliability and latency requirements. They formulated the autonomous resource allocation problem as the ratio of total rate to energy consumption, aiming to maximize the total energy efficiency of power-saving users under reliability and latency requirements. Since the energy efficiency problem is a complex mixed integer programming problem, a traffic-based resource allocation density heuristic algorithm was proposed to address this problem, ensuring the same successful transmission rate as perception-based algorithms while improving energy efficiency by reducing the power consumption per user [24].

Hegde et al. focused on the efficiency of radio resource allocation and scheduling algorithms in C-V2X communication networks and their impact on latency and reliability. Due to the continuous movement of vehicles, perception-based SPS becomes ineffective, leading to noneffective resource allocation and frequent resource conflicts. Therefore, the C-V2X communication network was described as a decentralized multi-agent network Markov decision process. Two variants, independent actor-critic and shared experience actor-critic, were proposed, achieving a 15–20% improvement in reception probability in

high vehicle density scenarios [25]. Saad et al. considered optimizing the medium access control layer in NR-V2X for more effective congestion control. They took into account the AoI indicator in the optimization process and introduced DRL to manage packet transmission rate and transmission power while ensuring high throughput. Compared with traditional distributed congestion control algorithms, the proposed solution demonstrated better performance in terms of timeliness, throughput, and average CBR. It highlights the importance and effectiveness of DRL-based congestion control mechanisms in the context of AoI [26].

Currently, NOMA has been applied in many scenarios. Ju et al. introduced an energy-efficient sub-channel and power allocation strategy for URLLC-enabled GF-NOMA systems using multi-agent Deep Reinforcement Learning (MADRL). It aims to maximize network energy efficiency while meeting URLLC requirements. It used simulation to check MA2DQN and MADQN performances [27]. Tran et al. explored secure offloading in vehicular edge computing (VEC) networks with malicious eavesdroppers using NOMA. An A3C-based scheme was proposed to optimize energy consumption and computation delay. Simulation results demonstrated its advantage in terms of energy efficiency and security [28]. Long et al. focused on VEC systems, where tasks can be processed locally or offloaded based on vehicle-to-infrastructure (V2I) and vehicle-to-vehicle (V2V) communication. It employed decentralized DRL and the deep deterministic policy gradient (DDPG) algorithm to perform the power allocation while addressing the uncertainty of MIMO-NOMA-based V2I communication and random task arrivals [29].

In summary, the existing works in the literature on NR-V2X performance analysis and research have not considered the impact of NOMA on AoI. They also have not considered employing DRL in optimizing AoI and energy consumption in NOMA and NR-V2X based vehicular networks. Therefore, we have undertaken the research presented in this paper.

3. System Model and Problem Formulation

3.1. Scenario Description

This section describes the AoI and energy optimization model based on NR-V2X Internet of Vehicles (IoV) system as shown in Figure 1, focusing primarily on the NR-V2X resource selection method. In Mode 2, vehicles utilizing NR-V2X adopt a perception-based SPS scheme for dynamic and semi-persistent resource selection. In the dynamic scheme, resources can only be used once, while in the semi-persistent scheme, resources are reserved for RC times. Additionally, the re-evaluation mechanism in Mode 2 can detect and avoid potential conflicts in message propagation. For NR-V2X Side-Link communication, resources in the time domain are composed of frames and subframes. Each frame typically consists of 10 subframes, with each frame's duration usually being 10 ms. The subframe is typically 1 ms. In the frequency domain, the smallest schedulable frequency unit is the Resource Block (RB). In the NR-V2X standard, RBs are sequentially combined to form subchannels, allowing vehicles to transmit messages on one or more subchannels [30]. Vehicles continuously monitor channels over a period of time by measuring the Reference Signal Received Power (RSRP) of all J subchannels, storing the latest information of N_{sense} time slots for use as a perception window when resource selection is required. RSRP represents the received power level of the reference signal in a mobile communication system, which is a key indicator for evaluating wireless signal quality and coverage. The higher the RSRP value, the stronger the received signal, and the better the signal quality. Then, vehicles initialize a selection window (SW) with a set of consecutive candidate time slots, namely RRI-sized slots. Each vehicle utilizes the information in the perception window to select available communication resources within the SW. Initially, Z_A is set to include all slots in the SW, and if z_n represents the time slot n following the perception window, then $Z_A = [z_n + T_1, z_n + T_1 + 1, \dots, z_n + \Gamma]$. Then, vehicles exclude all resources corresponding to the subframe from the set Z_A based on certain conditions. Firstly, due to half-duplex communication, vehicles cannot perceive the resources used by other vehicles in the same slot in the perception window. Hence, all resources corresponding to the slot

need to be excluded from the SW. Secondly, if the RSRP measurement corresponding to the candidate subframe exceeds $RSRP_{th}$, all resources corresponding to that candidate subframe are excluded from the SW. The exclusion criterion for the RSRP of the i -th subframe (the j -th subchannel) in the SW can be expressed as follows:

$$RSRP\left(z_{n+T1+i-N_{sense}}^j\right) \geq RSRP_{th}. \quad (1)$$

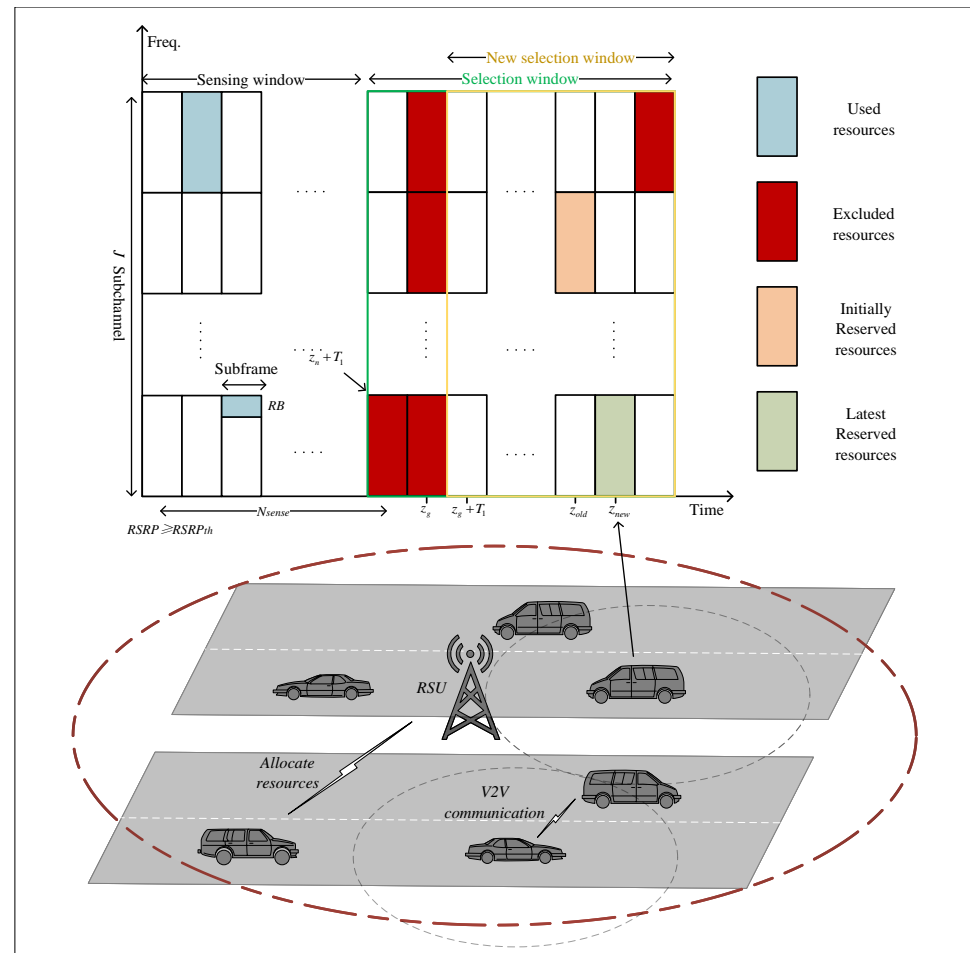


Figure 1. NR-V2X IoV system.

If the remaining number of resources in Z_A is less than $X\%$ of the total available resources, $RSRP_{th}$ is increased by 3 dB. In NR-V2X Mode 2, X can be set to 20, 35, or 50. Finally, vehicles randomly select a communication resource from the remaining resources in Z_A to reserve for subsequent transmission use, transmitting RC times at intervals of RRI. RC varies with RRI to ensure the time of the selected resource is between 0.5 and 1.5 s. Therefore, as shown in reference [31], the initial value of RC for a vehicle RC_0 is represented as follows:

$$RC_i^0 = \begin{cases} \frac{1000}{\max(20, \Gamma)} & 20 \leq \Gamma \leq 100 \\ 50 & 1 \leq \Gamma \leq 19 \end{cases}. \quad (2)$$

After RC decreases to 0, vehicles have the option to either continue using the pre-selected resources with a probability of P_{rk} or reselect new resources for transmission with a probability of $1 - P_{rk}$. Before transmitting messages, vehicles that select communication resources in time slot z_{old} may check whether these resources are still available using a reassessment mechanism (i.e., not reserved by another vehicle) [32]. Vehicles will perform

the reassessment check in time slot z_g . The new resource selection window, denoted as SW' , is defined as $[z_g + T_1, \dots, z_n + \Gamma]$. If resources previously excluded are found to be available again during the reselection process, vehicles will select new resources from the available resources in SW' . The resources initially chosen in time slot z_{old} will be replaced by new resources in time slot z_{new} as depicted in the figure. Table 1 lists the parameters used in this chapter.

Table 1. The summary for notations.

Notation	Description	Notation	Description
N_v	Total number of vehicles.	D	Road length.
D_{RSU}	RSU coverage range.	w	The maximum distance of the vehicle as the receiver.
RC	The value of the resource counter.	Γ	The size of SW and RRI.
η	SINR.	h_s	Small scale fading gain.
h	Large scale fading gain.	p	Vehicle transmission power.
L_d	Path loss.	d	Distance between communication vehicles.
I	Interference signal power.	p_n	noise power.
σ	Signal overlap degree.	C	Received signal power.
u	Communication situation.	W	Resource channel bandwidth for vehicle selection.
G	Message size.	Φ	The AoI at the receivers.
φ	The AoI of messages in the queue.	β	Has vehicle transmitted messages.
q	Queue length.	L	Queue capacity.
E	Energy consumption generated by vehicles in reserved resources.	l_i	Communication time of vehicle i .
N	Number of receivers for a certain vehicle.	$P^t(u^t = 1)$	The probability of the receiving end successfully receiving the message.
θ_x	The weights of the transition function strategy network between discrete and continuous actions.	θ_Q	The weights of state action value function strategy networks.
lr_x	Learning rate of transition function strategy networks between discrete and continuous actions.	lr_Q	Learning rate of state action value function strategy network.

3.2. NR-V2X-NOMA Communication Model

In the NR-V2X vehicular networking system, we denote the transmitter as i , receiver vehicle as j , and the considered time slot as t . SINR is represented as follows:

$$\eta_{i \rightarrow j}^t = \frac{h_s^t h_{i \rightarrow j}^t p_i^t / L_d(d_{i \rightarrow j})}{I_{i \rightarrow j}^t + p_n}, \quad (3)$$

where p_i^t is the transmission power by vehicle i , h_s^t is the random small-scale fading gain, $h_{i \rightarrow j}^t$ is the large-scale fading gain of the link from i to j in time slot t , $L(d_{i \rightarrow j})$ is the path loss as a function of the distance from i to j , p_n is the noise power, and $I_{i \rightarrow j}^t$ is the interference power. In Equation (3), the numerator represents the received power, while

the denominator is the sum of the noise power and interference, assumed to be Gaussian with zero mean. $I_{i \rightarrow j}^t$ is defined as follows:

$$I_{i \rightarrow j}^t = \sum_{k \in V^t, k \neq i} \sigma_{k,i}^t \frac{h_s^t h_{k \rightarrow j}^t p_k^t}{L_d(d_{k \rightarrow j})}, \quad (4)$$

where V^t is the set of nodes transmitting in time slot t , $\sigma_{k,i}^t$ is a multiplicative coefficient between 0 and 1 that quantifies the interference power of k in the subchannel used by i , relative to the transmission power of k . If k uses the exact same subchannel as i , $\sigma_{k,i}^t$ is 1; if its signal does not overlap or only partially overlaps, $\sigma_{k,i}^t$ is less than 1. The calculation of $\sigma_{k,i}^t$ takes into account the in-band emission, consistent with the specifications in [33].

The receiver j employs a serial interference cancellation mechanism in NOMA to decode multiple messages in different subchannels. The decoding process involves selecting the message with the maximum power as the desired signal, while treating the others as interference signals. If we denote the signal i with the current maximum received power as follows:

$$C_{i \rightarrow j}^t = \frac{h_s^t h_{i \rightarrow j}^t p_i^t}{L_d(d_{i \rightarrow j})}, \quad (5)$$

then, other signals with received power lower than signal i are denoted as follows:

$$C_{k \rightarrow j}^t = \frac{h_s^t h_{k \rightarrow j}^t p_k^t}{L_d(d_{k \rightarrow j})}. \quad (6)$$

The set of other vehicles whose received signal power is lower than vehicle i is denoted as $\mathcal{I}_i = \{k \in V_i \setminus i \mid C_k^t < C_i^t\}$. Then, the expression for the SINR obtained by vehicle j using NOMA is as follows:

$$\eta_{i \rightarrow j}^t = \frac{C_{i \rightarrow j}^t}{\sum_{k \in \mathcal{I}_i} C_{k \rightarrow j}^t \sigma_{k,i}^t + p_n}. \quad (7)$$

When vehicle j decodes the message with the highest power from vehicle i , the message power of the vehicle in \mathcal{I}_i is used as the interference power, and the magnitude of the interference is affected by the degree of channel overlap σ . Therefore, by adjusting the power allocation, the SINR of each message can be increased, thereby increasing the likelihood of successful communication.

3.3. AoI Model

When the size of the transmitted message is G , the criterion for successful communication between vehicles i and j is as follows:

$$u_{i \rightarrow j}^t = \left\lfloor \frac{W_i^t \log_2(1 + \eta_{i \rightarrow j}^t)}{G} \right\rfloor, \quad (8)$$

where $\lfloor \cdot \rfloor$ indicates rounding down the values in it, $\log_2(1 + \eta_{i \rightarrow j}^t)$ represents transmission rate, and W_i^t represents the bandwidth utilized by vehicle i for message transmission. So, $u_{i \rightarrow j}^t = 0$ indicates that the communication rate between vehicles i and j is insufficient to transmit the message within the specified time slot t , resulting in communication failure. Due to the nature of NR-V2X, where each transmission between vehicles requires waiting for a time equivalent to the RRI, each failed transmission between vehicles leads to an increase in the AoI at the receiving end vehicle. Successful transmission, on the other hand,

results in an increase in AoI at the receiving vehicle by Γ . The change in AoI at the receiving end vehicle j between communication time slots can be expressed as follows:

$$\Phi_{i \rightarrow j}^{t+\Gamma} = \begin{cases} \varphi_{in}^{t,1} + \Gamma & u_{i \rightarrow j}^t = 1 \\ \Phi_{i \rightarrow j}^t + \Gamma & u_{i \rightarrow j}^t = 0 \end{cases}. \quad (9)$$

It can be observed that the AoI at the receiving end is influenced by the transmission interval size Γ , transmission status u , and the AoI of the message transmitted by vehicle i , where $\Phi_{i \rightarrow i} = 0$. Among these factors, when the transmission interval Γ is smaller, the receiving end has more opportunities to update to the AoI of the transmission end. Additionally, a higher transmission success rate increases the likelihood of updating to the AoI at the transmission end. Furthermore, the AoI at the transmission end decreases as the queue processing rate $\frac{1}{\Gamma}$ increases. Therefore, when Γ is smaller, the AoI at the transmission end is smaller, resulting in a smaller AoI at the receiving end as well.

The average AoI at the receiving end for all vehicles in the system is defined as follows:

$$\bar{\Phi} = \frac{1}{T} \frac{1}{N_v^2} \sum_{i \in N_v} \sum_{j \in N_v} \Phi_{i \rightarrow j}, \quad (10)$$

In the scenario where four message types are considered, the $\varphi_{in}^{t+1,b}$ is represented as follows:

$$\varphi_{i,n}^{t+1,b} = \beta_{i,n}^t \left(\varphi_{i,n}^{t,b+1} - \varphi_{i,n}^{t,b} \right) + \varphi_{i,n}^{t,b} + 1, \quad (11)$$

In Equation (11), n represents multiple message types, b denotes the position of the message in the queue, and $\beta_{i,n}^t \in \{0, 1\}$ indicates whether a type n message can be transmitted (with 1 indicating it can be transmitted), where the queue operates on a first-in-first-out basis. And the parameter Γ determines the frequency of $\beta_{i,n}^t = 1$, so $\beta_{i,n}^t$ can be expressed as follows:

$$\beta_{i,n}^t = \left\lceil \frac{t}{z_{new} + m\Gamma} \right\rceil \left\lfloor \frac{q_{i,n}^t}{L} \right\rfloor. \quad (12)$$

where $\lceil \cdot \rceil$ represents rounding up the values in it, z_{new} represents the time slot allocated for vehicle reservation in the SW, m indicates the number of times a vehicle uses reserved resources, and q and L represent the queue length and queue capacity, respectively. Due to the consideration of multiple priority queues in this scenario, when the high-priority message is in queue $\beta_{i,n}^t = 1$, the $\beta_{i,n}^t = 1$ of other queues applies.

3.4. Energy Consumption Model

In NR-V2X, when vehicle i reselects resources, the energy consumption for the previously reserved resources is given by the following:

$$E_i^t = p_i^t l_i RC_i^0. \quad (13)$$

$p_i^t l_i$ represents the energy consumption generated by using reserved resources at a time, and l_i represents the time when vehicle i utilizes resources, which is the size of one time slot. As shown in Equation (2), when the transmission interval is smaller, RC_i^0 is larger, and the energy consumption during this time period is greater, indicating a trade-off between energy consumption and AoI.

The average energy consumption of all vehicles in the system is defined as

$$\bar{E} = \frac{1}{T} \frac{1}{N_v} \sum_{t \in \mathcal{T}} \sum_{i \in N_v} E_i^t. \quad (14)$$

4. Optimization Method of MPDQN Based on DRL

4.1. Framework for Optimization Problems

Based on the defined system model, the optimization problem is formulated to minimize the weighted sum of the average AoI and energy consumption of vehicles in the system. Since the AoI and energy consumption of vehicles depend on RRI Γ and power p , the optimization problem can be expressed as follows:

$$\min_{\Gamma^t, p^t} [\omega_1 \bar{E} + \omega_2 \bar{\Phi}] \quad (15)$$

$$s.t. \quad p^t \in [0, P_{max}], \forall t \in \mathcal{T}, \quad (15a)$$

$$\Gamma^t \in \{20, 50, 100\}, \forall t \in \mathcal{T}, \quad (15b)$$

where ω_1 and ω_2 are non-negative weight factors.

Since the channel conditions in the NR-V2X system are uncertain, we employ the Multi-Pass deep Q-Networks (MPDQN) method based on DRL to solve this optimization problem. In this method, the RSU serves as the agent, and its observed state at time slot t is as follows:

$$S^t = [s_1^t, \dots, s_i^t, \dots, s_{N_v}^t]. \quad (16)$$

The state of each vehicle is defined as follows:

$$s^t = [N^t, \bar{d}^t, P^t(u^t = 1), RC^0], \quad (17)$$

where N^t is the total number of other vehicles within the range w , defined as receivers. The average distance to these receivers is \bar{d}^t . $P^t(u^t = 1)$ is the probability of successful message reception by the receivers. RC^0 represents the total number of times that vehicles use reserved resources.

The action assigned by the RSU to vehicles at the time slot t is as follows:

$$a^t = (\Gamma^t, p_\Gamma), \quad (18)$$

where p_Γ represents the parameter that converts the continuous action p into a discrete action Γ . They act as two sub-actions, and the tuple they form constitutes the complete action assigned.

The objective of the optimization problem is to minimize the AoI and energy consumption in the system. Therefore, the reward function is defined as follows:

$$r_i^t = -(\omega_1 E_i^t + \omega_2 \bar{\Phi}_i^t). \quad (19)$$

where $\bar{\Phi}_i^t$ is the mean AoI of the receivers for vehicle i over a certain period of time:

$$\bar{\Phi}_i^t = \frac{1}{T} \frac{1}{N_v} \sum_{t=1}^T \sum_{j \in N_v} \Phi_{i \rightarrow j}^t. \quad (20)$$

4.2. Solution to Optimization Problems

For action tuple (Γ, p_Γ) , a policy network is used to match them.

$$p_\Gamma^t = x^Q(s^t, \Gamma; \theta_x), \quad (21)$$

where θ_x represents the weights of this network. Then, another deep neural network is used to approximate the action-value function $Q(s, (\Gamma, x))$:

$$Q(s^t, a^t) = Q\left(s^t, \left(\Gamma^t, x^Q(s^t, \Gamma; \theta_x)\right); \theta_Q\right), \quad (22)$$

where the network weights are denoted as θ_Q . The process where the agent obtains the action with the highest action value is illustrated in Figure 2.

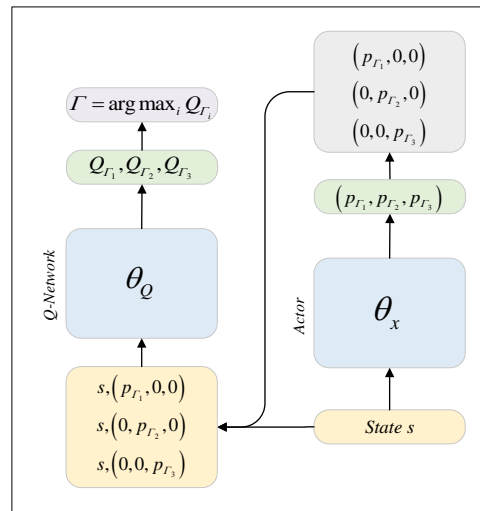


Figure 2. The process of agent selecting actions in MPDQN.

The loss functions for the Q-Network and the network of the actor are defined as follows:

$$L_Q(\theta_Q) = \mathbb{E}_{(s^t, (\Gamma, p_\Gamma), r^t, s^{t+1}) \sim M} \left[\frac{1}{2} (y^t - Q(s^t, (\Gamma, p_\Gamma); \theta_Q))^2 \right], \quad (23)$$

$$L_x(\theta_x) = \mathbb{E}_{s^t \sim M} \left[- \sum_{\Gamma \in \{20, 50, 10\}} Q(s^t, \Gamma, x^Q(s, \Gamma; \theta_x); \theta_Q) \right], \quad (24)$$

where y_t is defined as follows:

$$y^t = r^t + \gamma \max_{\Gamma} Q(s^{t+1}, \Gamma^{t+1}, x^Q(s^{t+1}; \theta_x); \theta_Q). \quad (25)$$

Finally, the network weights are updated using the learning rates lr_x and lr_Q for each network, aiming to approach the optimization objective.

$$\theta_Q^{t+1} = \theta_Q^t - lr_Q \nabla_{\theta_Q} L_Q(\theta_Q), \quad (26)$$

$$\theta_x^{t+1} = \theta_x^t - lr_x \nabla_{\theta_x} L_x(\theta_x), \quad (27)$$

Next, we will describe the algorithm process in detail. First, the parameters of both networks are randomly initialized, and an experience replay buffer of size M is established. Then, the algorithm iterates over EP episodes. At the beginning of each episode, the system parameters are reset. The RSU selects initial action tuples based on the initial state and the network and observes the next state. Subsequently, the algorithm iterates from timeslot 1 to timeslot T . For each timeslot t , the RSU allocates actions to vehicles needing resource reallocation based on the current state. When selecting actions, the RSU either explores randomly with a certain probability or chooses the action with the maximum Q-value, introducing exploration noise to avoid local optima. Finally, the tuple $[s^t, (\Gamma^t, p_\Gamma), r^t, s^{t+1}]$ is stored in the experience replay buffer. When the number of tuples in the buffer exceeds the sample size B , they will be employed to update the network parameters. The pseudocode is shown in Algorithm 1.

Algorithm 1: Optimization algorithm for AoI and energy consumption based on MPDQN

Input: γ, lr_x, lr_Q
Output: optimized θ_Q, θ_x
 Initialize the learning rates lr_x and lr_Q , experience replay buffer M , sample size B , network weights θ_x and θ_Q .
for episode from 1 to EP **do**
 Reset the model parameters;
 Observe the initial state s^0 and output the initial action (Γ_i^0, p_Γ) ;
 Obtain the state $s_i^1 = [N_i^1, d_i^1, Rn_i^1, RC_i^0]$.
 for slot t from 1 to T **do**
 Obtain action tuples based on the policy network or through random exploration:

$$a^t = \begin{cases} \text{random}(\Gamma^t, p_\Gamma) + \Delta & P_{ran} \\ \arg \max_{\Gamma} Q(s^t, \Gamma^t, x; \theta_Q) + \Delta & 1 - P_{ran} \end{cases}$$

 Execute action (Γ_i^t, p_Γ) , observe state s^{t+1} and reward r^t ;
 Store transition tuple $[s^t, (\Gamma^t, p_\Gamma), r^t, s^{t+1}]$ to M ;
 if number of tuples in M is larger than B **then**
 Randomly sample B transitions tuples from M ;
 Iterate networks according to Equation (23), (24), (26) and (27).
 end
 end
end

During the testing phase, there is no need to update the parameters. The actions are assigned to the required vehicles based on the optimized strategy in the training phase to carry out the test. The corresponding pseudocode is shown in Algorithm 2.

Algorithm 2: Testing stage of the MPDQN

for episode from 1 to EP **do**
 Reset the model parameters;
 Receive initial observation state s^1 ;
 for slot t from 1 to T **do**
 Assign RRI and power to vehicles that need to occupy communication resources according to the policy and their status.
 Execute action (Γ_i^t, p_Γ)
 Observe state s^{t+1} and reward r^t ;
 end
end

5. Simulation Results and Analysis

5.1. Parameter Settings

The simulation has been conducted using Python 3.6 and MATLAB 2023b, based on modifications and simulations built upon the code provided in [34]. The simulation scenario involves a two-way highway covered by the RSU communication range, where randomly distributed vehicles travel at constant speeds in their respective lanes and use NR-V2X Side-Link technology for V2V communication. The length of the highway, D , is 500 m, the RSU coverage range, D_{RSU} , is 250 m, and the maximum distance, w , between vehicles and receivers is 150 m. All vehicles utilize four different priority queues of length L . And the vehicles occupy a channel bandwidth of 10 MHz within the 5.9 GHz frequency band. The receiver has a noise figure of 9 dB. The path loss model features a standard deviation of 3 dB and a decorrelation distance of 25 m. Thus, the $RSRP_{th}$ is -126 dBm.

MPDQN employs a neural network with one hidden layer and updates its parameters using the Adam optimization method with learning rates $lr_Q = 5 \times 10^{-4}$ and $lr_x = 10^{-4}$. The experience replay buffer size M is set to 2000, and the sample size B is 128 [35]. Ornstein–Uhlenbeck noise is used as the exploration noise for the network, with a decay rate set to 0.15 and variance set to 0.0001. The key simulation parameters have been listed in Table 2.

Table 2. Values of the parameters in the experiments.

Parameter	Value	Parameter	Value
N_v	20, 30, 40, 50	L	10
D	500 m	D_{RSU}	250 m
w	150 m	P_{max}	23 dBm
v_{min}	60 km/h	v_{max}	80 km/h
W	10 MHz	$RSRP_{th}$	−126 dB
lr_Q	5×10^{-4}	lr_x	10^{-4}
τ	0.01	γ	0.99
M	2000	B	128

5.2. Simulation Result

In this section, we first compare the AoI of vehicles in LTE-V2X and NR-V2X. Then, we compare the AoI of vehicles in NR-V2X before and after using NOMA, as well as the AoI in LTE-V2X and NR-V2X based on NOMA. Finally, we optimize the joint objective of AoI and energy consumption in NR-V2X using MPDQN. Many recent works have employed genetic algorithms [36] and random algorithms [37] as baseline algorithms for resource allocation, and thus, we shall compare our approach with these two methods above.

Figure 3 illustrates the variation in average AoI in the system as the number of vehicles using LTE-V2X and NR-V2X for V2V direct communication scenarios. The number of vehicles considered are 20, 30, 40, and 50, with each vehicle following 3GPP standards and employing a random strategy to select its RRI (Resource Reuse Interval) and transmission power. It is observed that the average AoI in the system increases with the number of vehicles, regardless of whether the LTE-V2X or NR-V2X communication mode is utilized. This AoI increase can be attributed to the expansion of the receiver set within the communication range of vehicles as the number of vehicles in the system increases, leading to increased interference among them. Furthermore, due to the half-duplex communication mode, more vehicles are unable to receive messages due to resource contention, resulting in an increase in the average AoI. Additionally, as NR-V2X serves as a complement and advancement to LTE-V2X, the average AoI of vehicles in a vehicular networking system using NR-V2X for communication operations is consistently lower than that in an LTE-V2X system.

Figure 4 illustrates the variation in average AoI in the system as the number of vehicles changes when vehicles use NR-V2X and V2V communication with NOMA enabled. In this scenario, vehicles randomly select their RRI (Resource Reuse Interval) and transmission power. It can be observed that when vehicles use NOMA, they exhibit lower AoI when the number of vehicles changes. Moreover, the overall growth trend is smoother. This is attributed to NOMA's power-domain-based decoding approach, which significantly mitigates the impact of different vehicles occupying the same resources. Similarly, the average AoI within the system increases with the number of vehicles due to the proliferation of receivers. When more receivers cannot successfully receive messages due to resource contention, the AoI increases as the number of vehicles become larger.

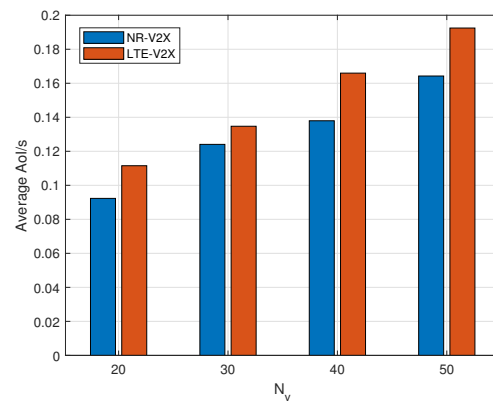


Figure 3. Average AoI between LTE-V2X and NR-V2X.

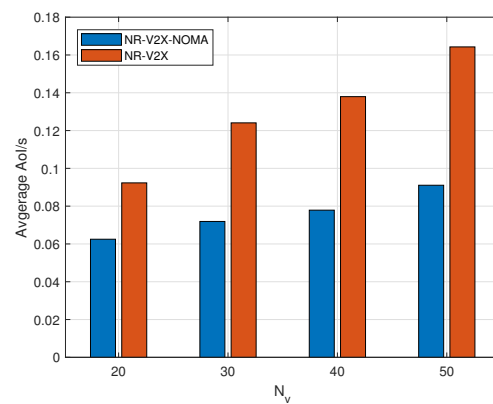


Figure 4. Average AoI before and after using NOMA for NR-V2X.

Figure 5 depicts the variation in average AoI within the system as the number of vehicles changes when vehicles utilize NOMA-based NR-V2X and LTE-V2X for V2V communication. With NOMA incorporated, the relationship between the two remains consistent with Figure 3, where the average AoI of NR-V2X consistently remains lower than that of LTE-V2X scenarios. This persistent advantage of NR-V2X over LTE-V2X can be attributed to the fact that vehicles in NR-V2X experience fewer resource collisions even before the implementation of NOMA.

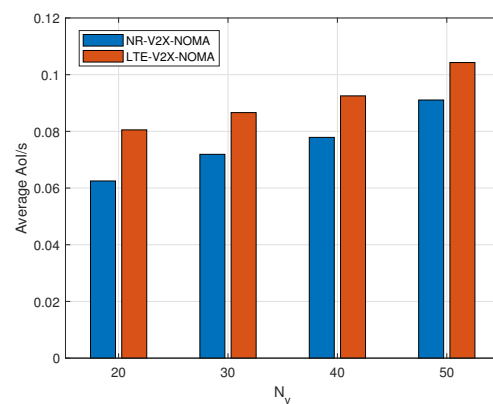


Figure 5. Average AoI between LTE-V2X and NR-V2X based on NOMA.

Figure 6 illustrates the learning curves of training under different scenarios. Overall, it can be observed that the rewards of different curves exhibit an upward trend in fluctuation from episode 0 to episode 500. Subsequently, the learning curves stabilize, indicating that

the agent has learned a strategy close to optimal. There is some jitter in the curves around episode 1000 when the number of vehicles is 40 and 50. This is attributed to detection noise impacting the agent, necessitating adjustments to return to a convergent state. Furthermore, it is noticeable that as the number of vehicles increases, the rewards decrease. This is due to the increasing interference experienced by each device with the growing number of devices in the system, resulting in lower SINR. This leads to prolonged transmission delays and increases the system AoI. To maintain a lower AoI, RSUs notify more vehicles to utilize communication resources for transmission, that is, more vehicles imply a higher AoI and increased energy consumption, and hence, lower rewards.

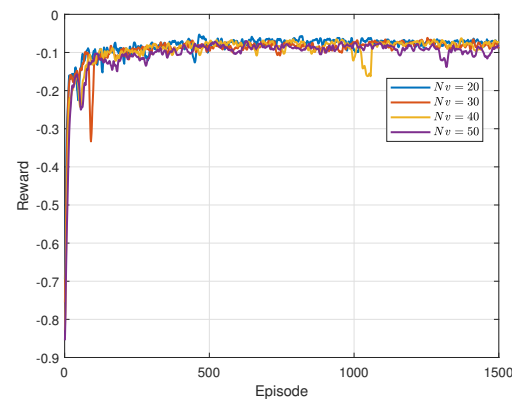


Figure 6. Learning curves under different numbers of vehicles.

Figure 7 depicts the variation in average AoI with the numbers of vehicles in the NOMA-based NR-V2X vehicular network system when employing MPDQN, genetic algorithm, and random algorithm strategies. It can be observed that the AoI of all three strategies increases with the increase in the number of devices. This is attributed to the interference experienced by each device as the number of devices increases, leading to increased transmission times according to the equation, which may further increase the AoI of systems. Furthermore, the allocation strategies obtained by MPDQN, which approximates the optimal strategy, and the genetic algorithm consistently outperform the random strategy. This is because the near-optimal strategy obtained by MPDQN selects actions for vehicles based on observed states, while the genetic algorithm derives better action allocation strategies through evolution, whereas the random strategy merely generates action allocations randomly. Additionally, it can be observed that the strategy derived from MPDQN outperforms that of the genetic algorithm, resulting in lower average AoI for vehicles using MPDQN. This is because MPDQN considers the impact of action allocations for each time slot on subsequent AoI, whereas the genetic algorithm does not.

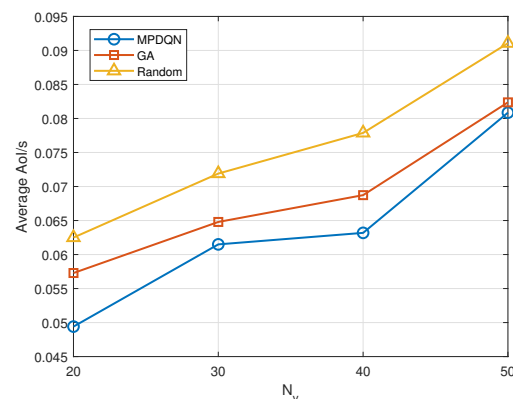


Figure 7. Average AoI with N_v under different algorithms.

Figure 8 compares the energy consumption within the system when vehicles employ three different methods. It can be observed that energy consumption increases when the number of devices increases. Moreover, the energy consumption in the random method does not exhibit significant variations for different numbers of vehicles within the system, resulting in a more linear energy consumption pattern. On the other hand, in MPDQN and genetic algorithm methods, the increase in the number of vehicles leads to an increase in interference power, resulting in a decrease in SINR. The reduced SINR leads to a longer transmission time and a higher probability of exceeding transmission time slots, thereby increasing the average information age of the system. However, due to the relatively larger weight of average energy consumption in the optimization objective, RSUs may choose to incur minimal additional energy consumption when the information age is low. Thus, the impact of the number of vehicles on energy consumption is less pronounced compared to its impact on information age. Furthermore, the strategies obtained by MPDQN and the genetic algorithm consistently outperform the random strategy, as they can derive better actions to ensure lower energy consumption costs at low AoI through their respective optimization approaches. Additionally, it can be observed that MPDQN consistently outperforms the genetic algorithm when the number of vehicles is high. This is because MPDQN has an advantage in handling real-time decision-making and dynamic environments, allowing it to continuously adjust strategies based on environmental feedback, resulting in stronger adaptability. In scenarios with a higher number of vehicles, MPDQN may have an easier time learning optimal scheduling strategies through interaction with the environment.

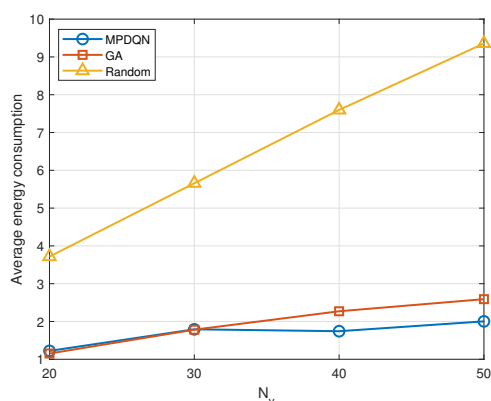


Figure 8. Average energy consumption with N_v under different algorithms.

Figure 9 compares the impact of different algorithms on average AoI in a scenario with 50 vehicles. It can be observed that as message size increases, average AoI also tends to increase. This is because the larger size of messages requires higher transmission rates, necessitating greater bandwidth and higher SINR. Among the three algorithms depicted above, MPDQN generally achieves the lowest average AoI, followed by the GA algorithm, highlighting the effectiveness of MPDQN.

Figure 10 compares the influence of different algorithms on average energy consumption in the same scenario as Figure 9. Here, energy consumption is averaged over the number of vehicles depicted in Figure 8. Consistent with the previous findings, MPDQN outperforms GA and random algorithms, ensuring lower energy consumption during vehicle communication in the scenario.

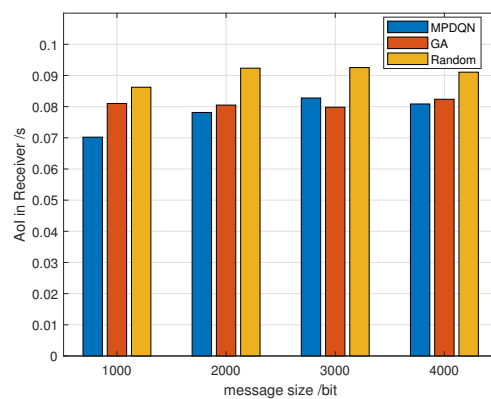


Figure 9. Average AoI with message size under different algorithms.

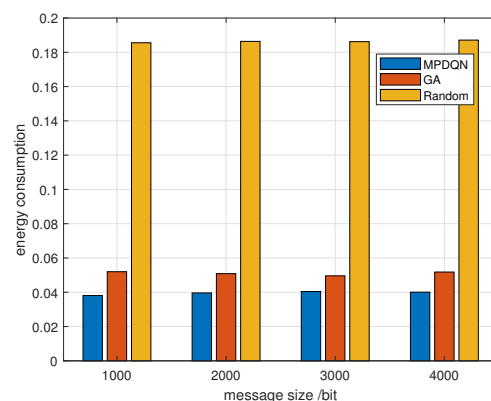


Figure 10. Average energy consumption with message size under different algorithms.

6. Conclusions

This paper addresses the probability of resource collisions in NR-V2X Mode 2 communication, despite the use of autonomous resource selection and probability reselection mechanisms. To mitigate the impact of collisions on the communication process, we proposed utilizing NOMA's serial interference cancellation mechanism. Additionally, we employed the MPDQN algorithm to dynamically adjust the transmission interval and transmission power of vehicles to reduce the average information age and energy consumption in the system. Firstly, we established communication models for NR-V2X and NOMA and then constructed a reinforcement learning framework based on MPDQN. In the learning framework, we modified the action space to enable simultaneous scheduling of discrete and continuous actions and finally optimized the joint problem of information age and energy consumption. Through simulation analysis, we demonstrated the advantages of NR-V2X over LTE-V2X, the improvement effect of NOMA on the information age performance in NR-V2X scenarios, and the effectiveness of MPDQN in reducing the information age and energy consumption in NR-V2X scenarios. Some potential challenges in this direction are noted: Future vehicles may use both LTE-V2X and NR-V2X for communication, so considering the coexistence and integration of these two technologies is a challenge [38]. In addition, fairness is often considered a key factor in NOMA-related scenarios [39]. Therefore, our future research will focus on performance optimization in scenarios where LTE-V2X and NR-V2X coexist and fairness in resource allocation in relevant scenarios with NOMA. In addition, MPDQN is a combination of DQN and DDPG, so the optimal performance can be improved by improving them intuitively.

Author Contributions: Conceptualization, S.S., Z.Z. and Q.W.; Methodology, S.S., Z.Z. and Q.W.; Software, S.S. and Z.Z.; Writing—Original Draft Preparation, Z.Z.; Writing—Review and Editing, P.F. and Q.F. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Natural Science Foundation of China under Grant No. 61701197, in part by the National Key Research and Development Program of China under Grant No. 2021YFA1000500(4), and in part by the 111 project under Grant No. B23008.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: Author Qiang Fan was employed by the company Qualcomm. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Wu, Q.; Wang, S.; Ge, H.; Fan, P.; Fan, Q.; Letaief, K.B. Delay-Sensitive Task Offloading in Vehicular Fog Computing-Assisted Platoons. *IEEE Trans. Netw. Serv. Manag.* **2024**, *21*, 2012–2026. [\[CrossRef\]](#)
2. Wu, Q.; Zheng, J. Performance modeling and analysis of the ADHOC MAC protocol for vehicular networks. *Wirel. Netw.* **2016**, *22*, 799–812. [\[CrossRef\]](#)
3. Chen, W.; Dai, L.; Letaief, K.B.; Cao, Z. A unified cross-layer framework for resource allocation in cooperative networks. *IEEE Trans. Wirel. Commun.* **2008**, *7*, 3000–3012. [\[CrossRef\]](#)
4. Zhang, Y.J.; Letaief, K.B. Adaptive resource allocation and scheduling for multiuser packet-based OFDM networks. In Proceedings of the 2004 IEEE International Conference on Communications (IEEE Cat. No. 04CH37577), Paris, France, 20–24 June 2004; IEEE: Piscataway, NJ, USA, 2004; Volume 5, pp. 2949–2953.
5. Wu, Q.; Wang, W.; Fan, P.; Fan, Q.; Wang, J.; Letaief, K.B. URLLC-Aware Resource Allocation for Heterogeneous Vehicular Edge Computing. *IEEE Trans. Veh. Technol.* **2024**, 1–16. [\[CrossRef\]](#)
6. Wu, Q.; Wang, X.; Fan, Q.; Fan, P.; Zhang, C.; Li, Z. High stable and accurate vehicle selection scheme based on federated edge learning in vehicular networks. *China Commun.* **2023**, *20*, 1–17. [\[CrossRef\]](#)
7. Jing, F.; Qiong, W.; JunFeng, H. Optimal deployment of wireless mesh sensor networks based on Delaunay triangulations. In Proceedings of the 2010 International Conference on Information, Networking and Automation (ICINA), Kunming, China, 17–19 October 2010; IEEE: Piscataway, NJ, USA, 2010; Volume 1, pp. V1-370–V1-374.
8. Qiong, W.; Shuai, S.; Ziyang, W.; Qiang, F.; Pingyi, F.; Cui, Z. Towards V2I age-aware fairness access: A DQN based intelligent vehicular node training and test method. *Chin. J. Electron.* **2023**, *32*, 1230–1244. [\[CrossRef\]](#)
9. Wu, Q.; Zhao, Y.; Fan, Q.; Fan, P.; Wang, J.; Zhang, C. Mobility-aware cooperative caching in vehicular edge computing based on asynchronous federated and deep reinforcement learning. *IEEE J. Sel. Top. Signal Process.* **2022**, *17*, 66–81. [\[CrossRef\]](#)
10. Wu, Q.; Wang, W.; Fan, P.; Fan, Q.; Zhu, H.; Letaief, K.B. Cooperative Edge Caching Based on Elastic Federated and Multi-Agent Deep Reinforcement Learning in Next-Generation Networks. *IEEE Trans. Netw. Serv. Manag.* **2024**, *1*. [\[CrossRef\]](#)
11. Tlake, L.C.; Markus, E.D.; Abu-Mahfouz, A.M. A Review of Interference Challenges on Integrated 5G NR and NB-IoT Networks. In Proceedings of the 2021 IEEE AFRICON, Arusha, Tanzania, 13–15 September 2021; pp. 1–6. [\[CrossRef\]](#)
12. Garcia, M.H.C.; Molina-Galan, A.; Boban, M.; Gozalvez, J.; Coll-Perales, B.; Şahin, T.; Kousaridas, A. A Tutorial on 5G NR V2X Communications. *IEEE Commun. Surv. Tutor.* **2021**, *23*, 1972–2026. [\[CrossRef\]](#)
13. Hu, J.; Yang, K.; Wen, G.; Hanzo, L. Integrated Data and Energy Communication Network: A Comprehensive Survey. *IEEE Commun. Surv. Tutor.* **2018**, *20*, 3169–3219. [\[CrossRef\]](#)
14. Wijerathna Basnayaka, C.M.; Jayakody, D.N.K.; Ponnimbaduge Perera, T.D.; Vidal Ribeiro, M. Age of Information in an URLLC-enabled Decode-and-Forward Wireless Communication System. In Proceedings of the 2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring), Virtual Event, 25–28 April 2021; pp. 1–6. [\[CrossRef\]](#)
15. Zhang, Y.; Zhou, Y.; Zhang, S.; Gui, G.; Adebisi, B.; Gacanin, H.; Sari, H. An Efficient Caching and Offloading Resource Allocation Strategy in Vehicular Social Networks. *IEEE Trans. Veh. Technol.* **2023**, *73*, 5690–5703. [\[CrossRef\]](#)
16. Rehman, A.; Valentini, R.; Cinque, E.; Di Marco, P.; Santucci, F. On the Impact of Multiple Access Interference in LTE-V2X and NR-V2X Sidelink Communications. *Sensors* **2023**, *23*, 4901. [\[CrossRef\]](#)
17. Anwar, W.; Franchi, N.; Fettweis, G. Physical layer evaluation of V2X communications technologies: 5G NR-V2X, LTE-V2X, IEEE 802.11 bd, and IEEE 802.11 p. In Proceedings of the 2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall), Honolulu, HI, USA, 22–25 September 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1–7.
18. Gong, J.; Chen, X.; Ma, X. Energy-age tradeoff in status update communication systems with retransmission. In Proceedings of the 2018 IEEE Global Communications Conference (GLOBECOM), Abu Dhabi, United Arab Emirates, 9–13 December 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1–6.
19. Gu, Y.; Chen, H.; Zhou, Y.; Li, Y.; Vucetic, B. Timely status update in Internet of Things monitoring systems: An age-energy tradeoff. *IEEE Internet Things J.* **2019**, *6*, 5324–5335. [\[CrossRef\]](#)
20. Yin, B.; Zhang, S.; Cheng, Y. Application-oriented scheduling for optimizing the age of correlated information: A deep-reinforcement-learning-based approach. *IEEE Internet Things J.* **2020**, *7*, 8748–8759. [\[CrossRef\]](#)
21. Liang, H.; Wang, L.; Zhang, Y.; Shan, H.; Shi, Z. Extending 5G NR V2X Mode 2 to Enable Integrated Sensing and Communication for Vehicular Networks. In Proceedings of the 2023 IEEE/CIC International Conference on Communications in China (ICCC Workshops), Dalian, China, 10–12 August 2023; IEEE: Piscataway, NJ, USA, 2023; pp. 1–6.

22. Xiaoqin, S.; Juanjuan, M.; Lei, L.; Tianchen, Z. Maximum-throughput sidelink resource allocation for NR-V2X networks with the energy-efficient CSI transmission. *IEEE Access* **2020**, *8*, 73164–73172. [[CrossRef](#)]
23. Molina-Galan, A.; Lusvarghi, L.; Coll-Perales, B.; Gozalvez, J.; Merani, M.L. On the Impact of Re-evaluation in 5G NR V2X Mode 2. *IEEE Trans. Veh. Technol.* **2024**, *73*, 2669–2683. [[CrossRef](#)]
24. Soleymani, D.M.; Ravichandran, L.; Gholami, M.R.; Del Galdo, G.; Harounabadi, M. Energy-efficient autonomous resource selection for power-saving users in NR V2X. In Proceedings of the 2021 IEEE 32nd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Helsinki, Finland, 13–16 September 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 972–978.
25. Hegde, A.; Song, R.; Festag, A. Radio resource allocation in 5G-NR V2X: A multi-agent actor-critic based approach. *IEEE Access* **2023**, *11*, 87225–87244. [[CrossRef](#)]
26. Saad, M.M.; Tariq, M.A.; Seo, J.; Ajmal, M.; Kim, D. Age-of-information aware intelligent MAC for congestion control in NR-V2X. In Proceedings of the 2023 Fourteenth International Conference on Ubiquitous and Future Networks (ICUFN), Paris, France, 4–7 July 2023; IEEE: Piscataway, NJ, USA, 2023; pp. 265–270.
27. Ju, Y.; Cao, Z.; Chen, Y.; Liu, L.; Pei, Q.; Mumtaz, S.; Dong, M. NOMA-Assisted Secure Offloading for Vehicular Edge Computing Networks With Asynchronous Deep Reinforcement Learning. *IEEE Trans. Intell. Transp. Syst.* **2024**, *25*, 2627–2640. [[CrossRef](#)]
28. Tran, D.D.; Sharma, S.K.; Ha, V.N.; Chatzinotas, S.; Woungang, I. Multi-Agent DRL Approach for Energy-Efficient Resource Allocation in URLLC-Enabled Grant-Free NOMA Systems. *IEEE Open J. Commun. Soc.* **2023**, *4*, 1470–1486. [[CrossRef](#)]
29. Long, D.; Wu, Q.; Fan, Q.; Fan, P.; Li, Z.; Fan, J. A power allocation scheme for MIMO-NOMA and D2D vehicular edge computing based on decentralized DRL. *Sensors* **2023**, *23*, 3449. [[CrossRef](#)]
30. Molina-Masegosa, R.; Gozalvez, J.; Sepulcre, M. Comparison of IEEE 802.11 p and LTE-V2X: An evaluation with periodic and aperiodic messages of constant and variable size. *IEEE Access* **2020**, *8*, 121526–121548. [[CrossRef](#)]
31. Ali, Z.; Lagén, S.; Giupponi, L.; Rouil, R. 3GPP NR V2X Mode 2: Overview, Models and System-Level Evaluation. *IEEE Access* **2021**, *9*, 89554–89579. [[CrossRef](#)] [[PubMed](#)]
32. Dayal, A.; Shah, V.K.; Dhillon, H.S.; Reed, J.H. Adaptive RRI Selection Algorithms for Improved Cooperative Awareness in Decentralized NR-V2X. *IEEE Access* **2023**, *11*, 134575–134588. [[CrossRef](#)]
33. Lin, X.; Li, J.; Baldemair, R.; Cheng, J.F.T.; Parkvall, S.; Larsson, D.C.; Koorapaty, H.; Frenne, M.; Falahati, S.; Grovlen, A.; et al. 5G new radio: Unveiling the essentials of the next generation wireless access technology. *IEEE Commun. Stand. Mag.* **2019**, *3*, 30–37. [[CrossRef](#)]
34. Todisco, V.; Bartoletti, S.; Campolo, C.; Molinaro, A.; Berthet, A.O.; Bazzi, A. Performance analysis of sidelink 5G-V2X mode 2 through an open-source simulator. *IEEE Access* **2021**, *9*, 145648–145661. [[CrossRef](#)]
35. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
36. Bulut, I.S.; Ilhan, H. Energy harvesting optimization of uplink-NOMA system for IoT networks based on channel capacity analysis using the water cycle algorithm. *IEEE Trans. Green Commun. Netw.* **2020**, *5*, 291–307. [[CrossRef](#)]
37. Ullah, S.A.; Zeb, S.; Mahmood, A.; Hassan, S.A.; Gidlund, M. Deep RL-assisted Energy Harvesting in CR-NOMA Communications for NextG IoT Networks. In Proceedings of the 2022 IEEE Globecom Workshops (GC Wkshps), Rio de Janeiro, Brazil, 4–8 December 2022; pp. 74–79. [[CrossRef](#)]
38. Naik, G.; Choudhury, B.; Park, J.M. IEEE 802.11 bd & 5G NR V2X: Evolution of radio access technologies for V2X communications. *IEEE Access* **2019**, *7*, 70169–70184.
39. Muhammed, A.J.; Ma, Z.; Diamantoulakis, P.D.; Li, L.; Karagiannidis, G.K. Energy-efficient resource allocation in multicarrier NOMA systems with fairness. *IEEE Trans. Commun.* **2019**, *67*, 8639–8654. [[CrossRef](#)]

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.