

Reinforcement learning-empowered resource allocation with multi-head attention mechanism in V2X networks

Irshad Khan, Manjula Sunkadakatte Haladappa

Department of Computer Science and Engineering, University of Visvesvaraya College of Engineering, Bengaluru, India

Article Info

Article history:

Received Apr 3, 2024

Revised Jul 8, 2024

Accepted Jul 16, 2024

Keywords:

Intelligent transport system
Multi-head attention mechanism
Reinforcement algorithm
Resource allocation
Vehicle to everything

ABSTRACT

Intelligent transport systems (ITS) offer safe and autonomous service in vehicular applications. The vehicle to everything (V2X) network aids in performing communication between any vehicle to other entities such as networks, pedestrians or other objects. However, the allocation of power in the V2X network is still seen as a challenging task in recent resource allocation approaches. So, multi-head attention mechanism with reinforcement learning (MHAMRL) is utilized in resource allocation. This work considers real traffic scenes in highway traffic model and wireless transmission model. Specifically, in the mode 4 cellular V2X, every individual vehicle is considered as a resource which does not rely on the base station for resource allocation. Vehicle users are classified into V2I or V2V links based on the varied service requirements of V2X. The combination of multi-head attention mechanism sequences the signal with minimal noises which diminishes the energy consumption and improves channel gain. In the velocity range of 20-25 m/s, the proposed approach achieves a sum rate of 53 Mb/s, surpassing the 50 Mb/s achieved by the existing multi-agent deep reinforcement learning-based attention mechanism (AMARL) algorithm.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Irshad Khan

Department of Computer Science and Engineering, University of Visvesvaraya College of Engineering
Bengaluru, India

Email: research.irshad@gmail.com

1. INTRODUCTION

In recent days, autonomous driving technology has witnessed rapid development in automotive network technology, helping in applications related to intelligent transport system (ITS). The vehicle-to-everything (V2X) communication is getting embodied with ITS due to the arising necessities of traffic management and transport requirements [1]–[3]. The V2X aids in performing communication among the network entities, servers and pedestrians. The V2X communication technology is facilitated with two primary radio standards, dedicated short range communication (DSRC) and cellular V2X (C-V2X) [4], [5]. When compared to DSRC, the C-V2X provides better convergence and quality of service (QoS) combined with advancements of non-orthogonal multiple access [6]. The V2X covers different kinds of vehicle communications such as vehicle-to-vehicle (V2V), Vehicle-to-Pedestrian (V2P), vehicle to infrastructure (V2I) and vehicle to network (V2N) [7]. Among these communication modes in V2X network, V2I and V2V communications are widely utilized in different vehicular applications. Entertainment and traffic related applications like streaming video and detection of crowd are monitored by vehicles with V2X communication technology [8]–[10].

The persistent packet collisions are a major drawback in V2X communication which only selects the nearby vehicles with limited frequency and delivers poor and inappropriate information [11], [12]. The

shortage or inappropriate radio sources among V2X network results in poor reliable and high latency communication which affects the overall process of allocating resources [13]–[15]. Thus, an effective resource allocation scheme has a significant role to play in enhancing communication performance among the platoon systems. The available resources are allocated along with the size of the platoon to enhance the capacity of the platoon and minimize energy consumption [16]. When the vehicles are in a congested formation, it is a matter of minimal time after which vehicular safety is threatened. V2X communication is relied upon within this critical time frame [17], [18]. When information is exchanged in the V2X network, every individual member of the platoon reacts to it and sends information about the presence of obstacles [19]. In external applications like V2X systems, there are numerous challenges for channel such as the managing environmental conditions which is complex [20]. Ensuring safety in V2X needs high-quality connection from vehicle to server to remain constant at vehicle movements that interrupt through obstacles such as trees, traffic vehicles, and buildings is to be considered [21]. The allocation of power in the V2X network is still seen as a challenging task in recent resource allocation approaches. This work considers the resource allocation problem prevailing in platoon-based vehicular network combined with V2I and V2V. When competitive relationship occurs, V2V link acts in an egoistic manner that disturbs communication efficiency. The multi-head attention mechanism is presented to exchange data among V2V links and each agent gains the ability to access relevant data.

The recent research on platoon-based resource allocation framework for C-V2X using various approaches carried out are discussed in this section. Han *et al.* [22] introduced a spectrum sharing scheme which was constrained by platoon based longitudinal control stability. Initially, time varying delay of communication system was analyzed for verifying the stability and pointing out errors that weaken the robustness of the platoon. Additionally, distribution resource allocation algorithm was introduced to optimize power allocation and spectrum matching among the vehicles in platoon. However, the suggested spectrum sharing scheme involved a single link pair that eliminated the reused spectrum. Sroka *et al.* [23] introduced a self-directed controller for scheduling awareness in platoon V2V communications. The suggested approach effectively minimized the usage of radio during the platoon maintenance stage but did not focus on string stability of the platoon which affected data transmission.

Zhao *et al.* [24] introduced a multi-agent deep reinforcement learning (MADRL) scheme for allocating resources in V2X networks. The MADRL scheme utilized state observations, action space and reward function to meet the QoS requirement of various users. However, the scalability which acted as a significant parameter in vehicular networks was not considered. Sabeeh *et al.* [25] introduced a centralized resource allocation scheme with spectrum re-partitioning technique for highway scenario. The spectrum re-partitioning approach along with frequency reuse techniques in roadside units (RSUs) were taken into consideration to avoid collisions with enhanced frequency reuse distance. However, the proposed approach was not capable of decoding packets from both transmitters due to the limited area coverage.

Parvini *et al.* [26] introduced a distributed resource allocation framework on the basis of multi-agent reinforcement learning (MARL), in which, every individual platoon leader acted as an agent and performed interaction in V2X networks. The suggested approach utilized holistic reward function which led to the enhancement of convergence rate. However, the distributed resource allocation framework did not focus on the scenarios of spectrum sharing. Ding *et al.* [27] introduced an effective framework known as multi-agent deep reinforcement learning-based attention mechanism (AMARL) to increase communication efficacy among V2X networks. In the suggested approach V2V link acted as an agent to perform intercommunication among the networks. Also, the attention mechanism was introduced to obtain appropriate information of vehicle but, the interaction among the vehicle and environment was not considered. Based on the above analysis, the existing techniques have several limitations such as spectrum sharing scheme involving a single link pair that eliminated the reused spectrum. Further, they did not focus on string stability of the platoon which affected data transmission. The scalability which acted as a significant parameter in vehicular networks was not considered. The models were not capable of decoding the packets from both transmitters due to limited area coverage and interaction among the vehicle and environment which was not considered. Moreover, distributed resource allocation framework did not focus on the scenarios of spectrum sharing. The major contributions of this work are listed as follows: i) The multi-head attention mechanism with reinforcement learning (MHAMRL) based resource allocation framework is proposed to solve the power allocation problem in V2X network; and ii) The V2I and V2V links are defined as agents based on resource allocation policy, observations, action space and reward function for MHAMRL resource allocation algorithm.

The remaining portions of this article are systematized as follows: section 2 outline of the proposed resource allocation framework. Section 3 discusses the results achieved through this approach. Section 4 presents the overall conclusions drawn from this study.

2. MULTI-OBJECTIVE RESOURCE ALLOCATION SCHEME

The multi-objective resource allocation scheme based on energy consumption rate and packet delivery rate is discussed. The proposed resource allocation framework considers two modes, namely, V2V and V2I. Additionally, the multi-objective V2X environment integrated with the state space, action space and reward function are explained in this section. A cross section of the road is considered for the scenario where the base station is placed at the center of the road. The vehicles communicate with each other via V2V link, whereas the V2I link is used for communication between vehicle and the base station. Each vehicle is equipped with a local server for its computation, and a centralized server is installed at the base station. The V2V and V2I communication modes in V2X networks are represented in Figure 1.

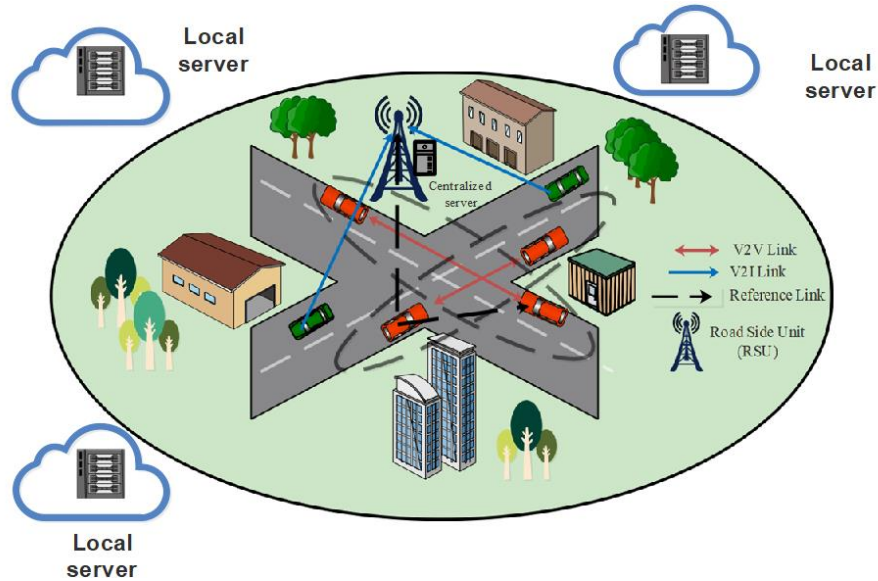


Figure 1. Modes of V2V and V2I communication takes place in V2X networks

2.1. System model

This work involves real traffic scenes in highway traffic model and wireless transmission model. Specifically in the mode 4 cellular V2X, every individual vehicle is considered an agent that does not rely on the base station for allocating resources. The users are classified according to V2I and V2V links as per the varied service requirements of V2X communications. When an interference takes place in the communication channel, it leads to disruption. The interference link helps to intimate information to the RSU and other vehicles in the communication channel. Furthermore, every individual V2V communicates only in one sub-band. The V2I and V2V communication links are assigned as $P = (1, \dots, P)$ and $Q = (1, \dots, Q)$, correspondingly. The total number of V2I and V2V links are represented as P and Q , respectively. The vehicles that are close to each other easily communicate amongst themselves but the vehicles that are far away perform communication with RSU to collect information. In this work, power gain is considered as the major concern while allocating resources and the channel gain is considered while transmitting packets from one channel to another channel in the V2X networks. The power gain of m^{th} V2I and n^{th} V2V links are denoted as $\hat{g}_{m,B}$ and $\hat{g}_{n,B}$, respectively. The interfering gain of the V2X channel over m^{th} V2I link from n^{th} V2V link over M^{th} sub-band is denoted as $g_{n,B}[M]$.

2.2. Signal interference in V2I link and V2V link

The signal interference is similar to the signal to noise ratio where the interference is specified to co-channel interference from other transmitting signals of V2X network. The signals received through the interference noise of m^{th} V2I link and n^{th} V2V link are represented in (1) and (2).

$$\varphi_m^{V2I}[m] = \frac{\hat{g}_{m,B} z_m^{V2I}}{\sum_{n=1}^N x_n[m] g_{n,B}[m] z_n^{V2V}[m] + \delta^2} \quad (1)$$

$$\varphi_n^{V2V}[m] = \frac{g_n[m] z_n^{V2V}[m]}{I_n[m] + \delta^2} \quad (2)$$

where, the power transmitted from m^{th} V2I link and n^{th} V2V link at m^{th} sub-band are represented as z_m^{V2I} and $z_n^{V2V}[m]$, respectively. The noise power occurred while communication is represented as δ^2 . The noise power is randomized in nature due to the varying amplitude and phases during the transmission of signals which is presented in (3).

$$I_n[m] = \hat{g}_{m,n} \cdot z_m^{V2I} + \sum_{n' \neq n}^N x_{n'}[m] \cdot g_{n',n}[m] \cdot z_{n'}^{V2V}[m] \quad (3)$$

where, $x_n[m]$ represents the spectrum allocation indicator which belongs to $\{0, 1\}$. If the value is 1, then it represents that the n^{th} V2V link uses the m^{th} sub band; or else, it is equal to 0. If V2V link accesses only one sub band, then $\sum_{m=1}^M x_n[m] \leq 1$ is satisfied. In this scenario, the efficiency of m^{th} V2I and n^{th} V2V links are expressed as shown in (4) and (5).

$$S_m^{V2I} = W \log(1 + \phi_m^{V2I}[m]) \quad (4)$$

$$S_n^{V2V}[m] = \sum_{m=1}^M x_n[m] \cdot W \log(1 + \phi_n^{V2V}[m]) \quad (5)$$

where, W represents the sub-band's bandwidth.

It is essential to optimize the bandwidth capacity of the V2I link for delivering high-quality entertainment services, satisfying the high reliability and low latency requirements of V2V link. This optimization ensures the provision of dependable data to the individual users. The primary criterion necessitates the maximization of the aggregate rate of V2I links. To fulfill the second criterion, V2V link must consistently transfer the size of packets within a fixed time T_{max} , adhering to the probabilistic model as shown in (6).

$$Pm \left\{ \sum_{t=1}^{T_{max}} S_n^{V2V}[m, t] \geq \frac{B}{\Delta_t} \right\} \quad (6)$$

where, Δ_t indicates the channel's coherence duration, and t is the index incorporated into $S_n^{V2V}[m, t]$ to represent how well the n^{th} V2V link performs across different slots during this coherence time. Therefore, the allocation issue of V2X resources is computed as an optimization task, as expressed in (7-9).

$$\max_{x, p^{V2V}} \sum_{m=1}^M S_m^{V2I} \quad (7)$$

$$s. t. \sum_{m=1}^M x_n[m] \leq 1 \quad (8)$$

$$z_n^{V2V}[m] \in P, \forall n, m \quad (9)$$

where, P signifies the discrete set of power levels for the V2V link. Equation (7) formulates the combinatorial optimization problem, constrained by the model's low accuracy requirement. Considering the dynamic nature of vehicle motion, the environment undergoes constant changes, introducing uncertainty in the parameters of the model. Obtaining and solving the whole complete channel state information (CSI) is difficult when conventional methods are employed.

2.3. Resource allocation based on reinforcement learning with multi-head attention mechanism

Following the stage of signal interference, multi-head attention with reinforcement learning is utilized to overcome the issues related to V2X resource allocation. Before introducing V2X resource allocation using the proposed approach, the functionalities involved in reinforcement algorithms are introduced. The functionalities of reinforcement algorithm are described in the following section.

2.3.1. Observation space

Because of vehicle motion, obtaining a comprehensive CSI is more difficult. Hence, partial channel state information (CSI) is included within the observation space, aligning more closely with the actual scenario. Simultaneously, this approach offers the advantage of minimizing overhead associated with the CSI response. In C-V2X mode 4, the vehicle assigns wireless resources based on channel measurements which inherently involves acquiring interference information, given the critical importance of minimizing delays in V2V networks. The V2V agent at time t encompasses the obtained interference details, residual time and payload. The observation space is numerically represented in (10).

$$observation = \{l_t^1, \dots, l_t^n, \dots, l_t^N\} \quad (10)$$

where, the set of agents are represented as n , observation of n^{th} agent at every time slot t is represented as l_t^n and is evaluated in (11).

$$l_t^n = \{\{l_{t-1}^n[m]\}_{m \in M}, p_t^n, r_t^n\} \quad (11)$$

where, the remaining payload and the remaining time are represented as p_t^n and r_t^n , respectively.

2.3.2. Action space

Considering the states observed, each individual V2V agent makes decisions by selecting the sub-band and allocating power. The action set of V2V agents is represented in (12).

$$a^t = \{a_n\}_{n=1}^N \quad (12)$$

where, the action space of n^{th} V2V agent is denoted as $a_n = \{s_n, p_n\}$. So, the set of action space while aligning the sub-bands for n^{th} V2V agent at time t is represented in (13).

$$S_t^n = \{s_t^n[1], \dots, s_t^n[m], \dots, s_t^n[m]\} \quad (13)$$

Likewise, the possible power selection set for the n^{th} V2V agent at time t is denoted in (14).

$$p_t^n \in \left\{0, \frac{1}{N-1} P_{max}, \frac{2}{N-1} P_{max}, \dots, P_{max}\right\} \quad (14)$$

where the set of probable sub-band pairs used to select the power among V2V agent is denoted as s_n, p_n , and the count of power levels are represented as N .

2.3.3. Rewards function

The rewards function is linked to the formulation in (7). The primary goal is to optimize throughput of V2I links which concurrently satisfies the reliability and latency necessities of V2V links. The reward function equation represented in (15), aiding in filling the latency requirements of V2V links.

$$R_t^n = \begin{cases} S_n^{V2V}(t), & p_t^n \geq 0 \\ c, & p_t^n < 0 \end{cases} \quad (15)$$

where, the hyper-parameter that is higher than maximum V2V link rate is represented as c . Transmission continues at an active rate of V2V link until the entire payload is successfully transferred. Furthermore, it is desirable to minimize the transmission time as much as possible, thereby increasing the possibility of effective transmission in a specified time. Hence, the final reward is represented as (16).

$$R_t = W_1 \sum_{m=1}^M S_m^{V2I}(t) + W_2 \sum_{n=1}^N R_t^n - W_3 (T_{max} - r_t^n) \quad (16)$$

Where, the weighted factor that represents the degree of different QoS is $\{W_i\}_{i=1,2,3}$.

2.4. Multi-head attention mechanism with reinforcement learning algorithm

The traditional attention mechanism acquires information about attention from a single agent. The multi-head attention conducts numerous linear transformation operations on the input feature matrix. This enables learning attention depiction under distinct linear conversions. Traditional attention techniques are computationally complex and minimal than the multi-head attention systems. Each self-attention layer is comprised with query matrix (Q_m), a key matrix (K_m), and a value matrix (V_m). As stated in (17), feature vector matrix T obtained from multichannel convolutional neural network, is an initial value for the query matrix (Q_m), key matrix (K_m), and value matrix (V_m).

$$Q_m = K_m = V_m = T_m \quad (17)$$

The scaled dot-product attention (SDA) is the central concept of the self-attention process. It calculates the attention scores by first taking the scaled dot product of Q_m and K_m , and then normalizing it by dividing by $\sqrt{d_k}$. This normalization ensures that the dot product result is appropriately scaled and prevents it

from becoming excessively large. The outcome is normalized based on softmax function and multiplied by matrix V_m to achieve the expression of attention. Equation (18) denotes the operation of SDA.

$$SDA(Q_m, K_m, V_m) = \text{Softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) V \quad (18)$$

where, the scaled dot product attention is represented as SDA and the dimensions of the K matrix are represented as d_k . The multi-head attention mechanism employs various parameters for performing linear transformations on matrices and linearly transforming the input. This results in scaled dot product attention, evaluated using (19).

$$\text{head}_i = SDA(Ql_i^Q, Kl_i^K, Vl_i^V) \quad (19)$$

where, the parameters used to perform linear transformations are represented as Ql_i^Q, Kl_i^K, Vl_i^V . The results are merged using the (20).

$$\text{Head} = \text{Multihead}(Q_m, K_m, V_m) = \text{merge}(\text{head}_1, \dots, \text{head}_h)l \quad (20)$$

where, the merged results are represented as head_1 and head_h . The parameter to complete the linear transformation is represented as l . The linear transformation is used to help the model adjust itself to varying tokens in the sequence based on their relationship with each other. Each agent autonomously engages with its surroundings to collect local observations and acquire information from others through multi-head attention mechanism for spectrum allocation. It is accomplished by altering the result matrix to remove the Head dimension and the steps involved in merging are as follows:

- The head and sequence dimensions in the attention score matrix is rearranged. In other words, the matrix shape transitions occur from (Batch B , Head H , Sequence S , Query size Q) to B, S, H and Q . The head dimension is moved to the third position because the attention heads are applied independently across different positions, in turn minimizing the complexity of the model while allocating resources.
- The head dimension is reshaped to B, S, H^* and Q . This basically concatenates each head's Attention Score vectors into a single attention. An approach with multi-head attention is implemented to address the optimization of the resource allocation problem outlined in (7). In this framework, each agent possesses its Q-network and channel gain. The multi-head attention mechanism with V2V links integrate the useful information into an evaluation of the estimation function. The overall process involved in the proposed approach is summarized in algorithm 1.

In each step's time, V2V combines the observed states and selects the better link to allocate resources with the optimal power gain and better throughput. When the resource is allocated among optimal power gain and better throughput, the terminal state is achieved.

Algorithm 1. MHAMRL

```

1 Input: V2X environment simulator, multi-head attention network, payload size and latency
2 Output: Weight of MHAMRL network
3 Set constraint for learning agents  $Q_m$  and  $O^T$ 
4 Obtain preliminary phase  $K_i^n$  of each agent (13)
5 for time = 1, ..., t do
6 Make action  $a^t$  based on Reinforcement learning network policy  $a_n = \{s_n, p_n\}$  (11)
7 Action is transferred to reward aggregator (16).
8 Achieve global reward  $V^t$  based on reward aggregator.
9 Detect new state (action space)  $O^{t+1}$ 
10 if terminal phase then
11 Aggregate actor gradient Q
12 Aggregate critic gradient  $R_i^n$  (15)
13 Update parameters  $R_i^n$  and  $Q$  based on (15) and (18)
14 end if
15 end for

```

3. RESULTS AND ANALYSIS

The results obtained from the proposed multi-head attention mechanism with reinforcement algorithm used in the process of allocating resources in V2X networks exhibit superior throughput and energy efficiency, as presented in this section. The scenario considered is a cross section of abroad with the base station being at the center having four lanes described as the *up_lane*, *down_lane*, *left_uplane*, and

right_downlane. Each vehicle is considered an agent. When all agents work together to share the resources, it is considered the multi-head agent. The V2X communication scenarios are created using Python and the system is equipped with an Intel i7 processor, 8 GB RAM and windows 11 operating system. The simulation parameters are listed in Table 1.

Table 1. Simulation parameter add other related parameters

Parameters	Values
Carrier frequency	2 GHz
Speed of the vehicle	10 m/s – 25 m/s
Noise power	-114 dBm
Latency time limit	24 ms

3.1. Performance analysis

The performance of the proposed approach is evaluated based on the sum rate of V2I and V2V links. Table 2 demonstrates the sum rate of V2V and V2I links for different power levels of 23, 25, 27, and 29 dBm for payload of 1×1060 , 2×1060 , 3×1060 , 4×1060 and 5×1060 bytes. The sum rate of V2I and V2V links for various power values in dBm are represented in Figures 2 and 3.

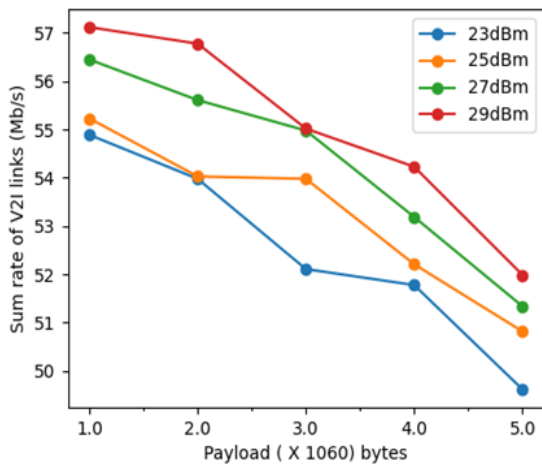


Figure 2. Sum rate of V2I link for various power values in dBm

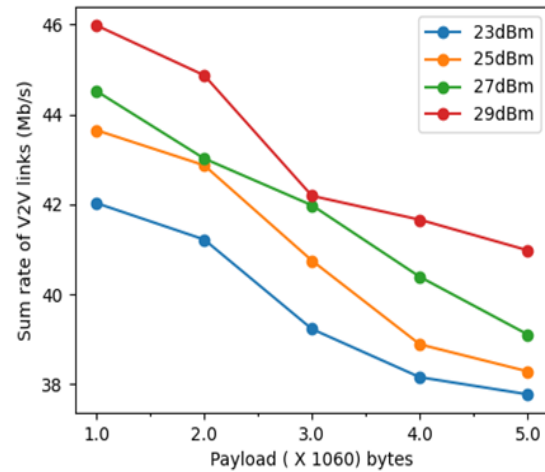


Figure 3. Sum rate of V2V link for various power values in dBm

When the payload value gets increased, the performance at all set power transmission speeds is gradually reduced, which is graphically represented in Figures 2 and 3. This is due to the fact that the payload rises; the links demand higher transmission periods and transmission power. This, in turn, induces higher interference within V2I and V2V links, leading to a reduction in communication efficiency. However, the proposed approach exhibits minimum variation at the time of transmission for varying payload values. Secondly, the performance of the proposed multi-head attention approach is also evaluated based on varying velocity of the vehicles for fixed payload and power value. Table 2 displays the impact of the velocity on sum rate of V2V link for resource allocation using the existing reinforcement algorithm and reinforcement algorithm with multi-head attention mechanism.

Table 2. Performance analysis for various velocity of V2V

Velocity of Vehicle (m/s)	Sum rate of V2V links(Mb/s)	
	Resource allocation using reinforcement algorithm	Resource allocation using reinforcement algorithm with multi-head attention mechanism
10-15	55	59
15-20	43	48
20-25	38	45

In a similar way, the performance of the proposed approach is evaluated based on the velocity of the vehicles and sum rate in the V2I link. The sum rate of V2I links with resource allocation with the help of reinforcement algorithm and resource allocation using the reinforcement algorithm with multi-head attention mechanism for different velocities is represented in Table 3. The results from Table 2 and Table 3 show that the proposed resource allocation scheme with the help of reinforcement algorithm and multi-head attention mechanisms achieves a better performance with respect to conventional resource allocation algorithm.

Table 3. Performance analysis for various velocity of V2I from 10 m/s to 25 m/s

Velocity of vehicle (m/s)	Sum rate of V2I links (Mb/s)	
	Resource allocation using reinforcement algorithm	Resource allocation using reinforcement algorithm with multi-head attention mechanism
10-15	54	67
15-20	51	59
20-25	48	51

For instance, when the velocity of the vehicle is 10m/s to 15 m/s, the sum rate of V2V links of the proposed approach is 59 Mb/s; whereas the conventional allocation scheme using reinforcement algorithm is 55 Mb/s. Similarly, for the velocity of 10-15 m/s, the sum of V2I links for resource allocation using the reinforcement algorithm is 54 Mb/s, and resource allocation with multi-head reinforcement algorithm is 67 Mb/s. The combination of multi-head attention mechanism sequences the signal with minimal noises. The signals with minimal noises diminish energy consumption and improve the channel gain.

3.2. Comparative analysis

In this section, the performance of the proposed reinforcement algorithm with multi-head attention mechanism (MHAMRL) is compared with the existing MADRL [24] and AMARL [27]. The comparison is performed based on the velocity of the vehicle ranges from 10 to 60 m/s. In Table 4, the comparison is performed based on the average V2V link's success probability. The comparison is performed with two different velocities 15 and 60 m/s.

The results from Table 4 show that the proposed MHAMRL achieves a superior average V2V link success probability. The average V2V link success probability of the proposed MHAMRL for velocity of 15 m/s is 0.998, whereas the existing MADRL obtains 0.993. The better success probability of the proposed approach helps in reliable data packet transmission with minimal noises. When the noises are diminished during transmission, the energy consumption is reduced with an improved channel gain along with increased success probability. In Table 5, the comparison is performed based on the sum rate of V2I links of the proposed approach, and AMARL for vehicle velocities ranges from 10-25 m/s for the payload value of 2×1060 bytes. The results from Table 5 prove the efficiency of the suggested methodology based on the summing rate of V2I links. The sum rate of the proposed method for the velocity of 20-25 m/s is 53 Mb/s, whereas the existing AMARL obtains a sum rate of 50Mb/s. Thus, the overall results from Tables 4 and 5 prove that the proposed approach achieves better summing capacity and success probability during resource allocation in V2X networks. Utilizing a multi-head attention mechanism effectively organizes the signal, ensuring minimal interference. The signals with reduced noise levels contribute to lower energy consumption and improved channel gain.

Table 4. Comparison of average V2V link success probability

Velocity of vehicle (m/s)	Average V2V link success probability	
	MADRL [24]	MHAMRL
15	0.993	0.998
60	0.918	0.925

Table 5. Sum rate comparison of V2I links

Velocity of vehicle (m/s)	Sum rate of V2I links (Mb/s)	
	AMARL [27]	MHAMRL
10-15	54	59
15-20	52	55
20-25	50	53

3.3. Discussion

The advantages and limitations of the proposed method of existing methods are discussed in this section. The existing methods such as the MADRL's [24] scalability acts as a significant parameter in vehicular networks was not considered. In AMARL [27], the interaction between vehicle and environment was not considered. The MHAMRL based resource allocation framework is proposed to solve the power allocation problem in V2X network. The V2I and V2V links are defined as agents based on resource allocation policy, observations, action space and reward function for MHAMRL resource allocation algorithm.

4. CONCLUSION

This work introduces multi-head attention mechanism based on reinforcement algorithm for optimizing sum rate and allocation of power in V2I and V2V links. The multi-head attention mechanism helps to perform an effective communication among V2V links and adapt to the varying environmental changes. The multi-head attention mechanism performs multiple linear transformations on the input feature matrix, acquiring attention representations under various linear transformations. The reinforcement learning combined with multi-head attention mechanism explores power variation of V2V and V2I networks and allocates the resources in an effective manner. The average V2V link's success probability of the proposed MHAMRL for velocity of 15 m/s is 0.998, while that of the existing MADRL is 0.993. These results exhibit the robustness of the proposed approach. However, the proposed approach does not consider the V2P and V2N models which can be considered to develop an effective model in future research.





REFERENCES

- [1] T. T. Thanh Le and S. Moh, "Comprehensive survey of radio resource allocation schemes for 5G V2X communications," *IEEE Access*, vol. 9, pp. 123117–123133, 2021, doi: 10.1109/access.2021.3109894.
- [2] F. Jameel, W. U. Khan, N. Kumar, and R. Jantti, "Efficient power-splitting and resource allocation for cellular V2X communications," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 6, pp. 3547–3556, Jun. 2021, doi: 10.1109/tits.2020.3001682.
- [3] Y. Jeon and H. Kim, "An explicit reservation-augmented resource allocation scheme for C-V2X sidelink mode 4," *IEEE Access*, vol. 8, pp. 147241–147255, 2020, doi: 10.1109/access.2020.3015549.
- [4] B. Gu, W. Chen, M. Alazab, X. Tan, and M. Guizani, "Multiagent reinforcement learning-based semi-persistent scheduling scheme in C-V2X mode 4," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 11, pp. 12044–12056, Nov. 2022, doi: 10.1109/tvt.2022.3189019.
- [5] C. Shin, E. Farag, H. Ryu, M. Zhou, and Y. Kim, "Vehicle-to-everything (V2X) evolution from 4G to 5G in 3GPP: Focusing on resource allocation aspects," *IEEE Access*, vol. 11, pp. 18689–18703, 2023, doi: 10.1109/access.2023.3247127.
- [6] X. Zhang, M. Peng, S. Yan, and Y. Sun, "Deep-reinforcement-learning-based mode selection and resource allocation for cellular V2X communications," *IEEE Internet of Things Journal*, vol. 7, no. 7, pp. 6380–6391, Jul. 2020, doi: 10.1109/jiot.2019.2962715.
- [7] T. I. Bayu, Y.-F. Huang, and J.-K. Chen, "Performance of fuzzy inference system for adaptive resource allocation in C-V2X networks," *Electronics*, vol. 11, no. 23, p. 4063, Dec. 2022, doi: 10.3390/electronics11234063.
- [8] Y. Yuan, G. Zheng, K.-K. Wong, and K. B. Letaief, "Meta-reinforcement learning based resource allocation for dynamic V2X communications," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 9, pp. 8964–8977, Sep. 2021, doi: 10.1109/tvt.2021.3098854.
- [9] D. Li, S. Xu, and P. Li, "Deep reinforcement learning-empowered resource allocation for mobile edge computing in cellular V2X Networks," *Sensors*, vol. 21, no. 2, p. 372, Jan. 2021, doi: 10.3390/s21020372.
- [10] A. Hegde, R. Song, and A. Festag, "Radio resource allocation in 5G-NR V2X: A multi-agent actor-critic based approach," *IEEE Access*, vol. 11, pp. 87225–87244, 2023, doi: 10.1109/access.2023.3305267.
- [11] K. Park and H. Kim, "Communication outages mitigation through mutual assistance for cellular V2X-based platooning," *2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring)*, Helsinki, Finland, 2022, pp. 1-5, doi: 10.1109/vtc2022-spring54318.2022.9860768.
- [12] Z. Dong et al., "Dynamic manager selection assisted resource allocation in URLLC with finite block length for 5G-V2X platoons," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 11, pp. 11336–11350, Nov. 2022, doi: 10.1109/tvt.2022.3190518.
- [13] Y. Wang, L. Zhang, C. Wei, and Y. Tang, "Joint optimization of resource allocation and computation offloading based on game coalition in C-V2X," *Ad Hoc Networks*, vol. 150, p. 103266, Nov. 2023, doi: 10.1016/j.adhoc.2023.103266.
- [14] N. Banitalebi, P. Azmi, N. Mokari, A. H. Arani, and H. Yanikomeroglu, "Distributed learning-based resource allocation for self-organizing C-V2X communication in cellular networks," *IEEE Open Journal of the Communications Society*, vol. 3, pp. 1719–1736, 2022, doi: 10.1109/ojcoms.2022.3211340.
- [15] K.-L. Chen, W.-Y. Chen, and R.-H. Hwang, "Optimizing resource allocation with high-reliability constraint for multicasting automotive messages in 5G NR C-V2X networks," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 4, pp. 4792–4804, Apr. 2023, doi: 10.1109/tvt.2022.3226215.
- [16] D. Han and J. So, "Energy-efficient resource allocation based on deep Q-network in V2V communications," *Sensors*, vol. 23, no. 3, p. 1295, Jan. 2023, doi: 10.3390/s23031295.
- [17] I. Lee and D. K. Kim, "Decentralized multi-agent dqn-based resource allocation for heterogeneous traffic in V2X communications," *IEEE Access*, vol. 12, pp. 3070–3084, 2024, doi: 10.1109/access.2023.3349350.
- [18] A. Bazzi, C. Campolo, A. Molinaro, A. O. Berthet, B. M. Masini, and A. Zanella, "On wireless blind spots in the C-V2X sidelink," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 8, pp. 9239–9243, Aug. 2020, doi: 10.1109/tvt.2020.3001074.
- [19] Z. Mlika and S. Cherkaoui, "Deep deterministic policy gradient to minimize the age of information in cellular V2X communications," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 12, pp. 23597–23612, Dec. 2022, doi: 10.1109/tits.2022.3190799.





- [20] A. Kharbouche, Z. Madini, and Y. Zouine, "Performance improvements of a VLC system, in a V2X context, using a different multiplexing technique," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 21, no. 4, p. 725, Aug. 2023, doi: 10.12928/telkomnika.v21i4.24042.
- [21] H.-N. Nguyen, N.-T. Nguyen, and G.-T. Vo, "On performance analysis of non-orthogonal multiple access downlink for cellular-connected unmanned aerial vehicle relaying assisted vehicle-to-everything system," *International Journal of Electrical and Computer Engineering*, vol. 14, no. 2, pp. 1634–1645, Apr. 2024, doi: 10.11591/ijece.v14i2.pp1634-1645.
- [22] Q. Han, C. Liu, H. Yang, and Z. Zuo, "Longitudinal control-oriented spectrum sharing based on C-V2X for vehicle platoons," *IEEE Systems Journal*, vol. 17, no. 1, pp. 1125–1136, Mar. 2023, doi: 10.1109/jsyst.2022.3201816.
- [23] E. Sroka Paweland Ström, T. Svensson, and A. Kliks, "Autonomous controller-aware scheduling of intra-platoon V2V communications," *Sensors*, vol. 23, no. 1, p. 60, Dec. 2022, doi: 10.3390/s23010060.
- [24] J. Zhao, F. Hu, J. Li, and Y. Nie, "Multi-agent deep reinforcement learning based resource management in heterogeneous V2X networks," *Digital Communications and Networks*, Jun. 2023, doi: 10.1016/j.dcan.2023.06.003.
- [25] S. Sabeeh, K. Wesołowski, and P. Sroka, "C-V2X centralized resource allocation with spectrum re-partitioning in highway scenario," *Electronics*, vol. 11, no. 2, p. 279, Jan. 2022, doi: 10.3390/electronics11020279.
- [26] M. Parvini, M. R. Javan, N. Mokari, B. Abbasi, and E. A. Jorswieck, "AoI-aware resource allocation for platoon-based C-V2X networks via multi-agent multi-task reinforcement learning," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 8, pp. 9880–9896, Aug. 2023, doi: 10.1109/tvt.2023.3259688.
- [27] Y. Ding, Y. Huang, L. Tang, X. Qin, and Z. Jia, "Resource allocation in V2X communications based on multi-agent reinforcement learning with attention mechanism," *Mathematics*, vol. 10, no. 19, p. 3415, Sep. 2022, doi: 10.3390/math10193415.

BIOGRAPHIES OF AUTHORS



Irshad Khan     is currently a full-time research scholar at the Department of Computer Science and Engineering, University of Visvesvaraya College of Engineering, Bangalore University, Bengaluru, India. He earned his bachelor's degree in computer science and engineering from Visvesvaraya Technological University and pursued a master's degree in computer networks from the same university. Irshad's expertise spans diverse domains, including computer networks, V2X communication, intelligent transportation systems, artificial intelligence, machine learning, neural networks, reinforcement learning, and autonomous vehicles. His dedication to cutting-edge research is evident through the presentation of two IEEE conference papers, underscoring his passion for pushing the boundaries of technological understanding and contributing to the future of computer science and intelligent transportation systems. He can be contacted at email: research.irshad@gmail.com.



Manjula Sunkadakatte Haladappa     currently a professor in the Department of Computer Science and Engineering at the University Visvesvaraya College of Engineering, Bangalore University, Bengaluru, has established herself as a prominent figure in the field. Holding a B.E., M.Tech., and Ph.D. in computer science and engineering, she has honed her expertise in computer networks, wireless sensor networks, data mining, cloud computing, artificial intelligence, machine learning, and federated learning. Driven by a passion for knowledge, she has made significant contributions to these domains, evident through her scholarly publications and conference presentations. As an educator, her influence extends beyond the classroom, shaping the minds of future engineers. Manjula's holistic approach to academia, balancing research and teaching, underscores her commitment to advancing the frontiers of computer science and engineering, leaving an indelible mark on both her students and the technological landscape at large. With an impressive track record, she has authored 69 journals, presented 72 conference papers, holds 5 patents, and has contributed to 4 book chapters while publishing 4 books, showcasing her prolific and diverse contributions to the academic and technological community. Additionally, Dr. Manjula is currently a respected member of the executive council and has served as a former member of the academic senate at VTU Belagavi. She can be contacted at email: shmanjula@gmail.com.