This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3294279

FAQ: A Fuzzy-Logic-Assisted Q Learning Model for Resource Allocation in 6G V2X

Minglong Zhang, Yi Dou, Vuk Marojevic, Peter Han Joo Chong and Henry C. B. Chan

***Abstract*—This research proposes a dynamic resource allocation method for vehicle-to-everything (V2X) communications in the six generation (6G) cellular networks. Cellular V2X (C-V2X) communications empower advanced applications but at the same time bring unprecedented challenges in how to fully utilize the limited physical-layer resources, given the fact that most of the applications require both ultra low latency, high data rate and high reliability. Resource allocation plays a pivotal role to satisfy such requirements as well as guarantee quality of service (QoS). Based on this observation, a novel fuzzy-logic-assisted Q learning model (FAQ) is proposed to intelligently and dynamically allocate resources by taking advantage of the centralized allocation mode. The proposed FAQ model reuses the resources to maximize the network throughput while minimizing the interference caused by concurrent transmissions. The fuzzy-logic module expedites the learning and improves the performance of the Q-learning. A mathematical model is developed to analyze the network throughput considering the interference. To evaluate the performance, a system model for V2X communications is built for urban areas, where various V2X services are deployed in the network. Simulation results show that the proposed FAQ algorithm can significantly outperform deep reinforcement learning, Q learning and other advanced allocation strategies regarding the convergence speed and the network throughput.**

***Index Terms*—Reinforcement Learning, 6G V2X, Resource Allocation, Fuzzy Logic, Vehicular Networks.**

I. Introduction

Cellular vehicular-to-everything communications (C-V2X) are specifically for connecting vehicles and vehicles, infrastructure and other smart user equipment (UE) to acquire safety, traffic and surrounding environment information. These information are used to improve road safety, transportation efficiency and even driving/riding comfort. To this end, four enhanced V2X (eV2X) services (e.g., extended sensing, vehicle platooning, as well as advanced and automated driving) and many basic V2X cases have been introduced in the fifth generation (5G) standard [1] [2].

Compared with the V2X communications in long-term evolution (LTE) [3], 5G V2X achieves better performance through adding more spectral and hardware resources but still inheriting certain underlying mechanisms and architectures. However, with the rapid increase of autonomous vehicles in the near future, explosive data communications will be demanded by various digital devices and applications. Meanwhile, many emerging services in autonomous vehicles, such as 3-D displays, holographic control display systems and immersive entertainment, will bring new scientific and technical challenges to the existing vehicular communications regarding data rate, coverage, latency, and intelligence. Unfortunately, 5G V2X may not well deal with the challenging situations. To address the limitations of 5G in this domain, the recent proposed 6G cellular communications [4], which aim to integrate terrestrial and several non-terrestrial communication networks, introduce several technical advancements, such as artificial intelligence (AI) and machine learning (ML), Terahertz (THz) communication, unmanned aerial vehicles (UAVs) and intelligent reflecting surface (IRS) [5], [6]. These will facilitate intelligent and ubiquitous C-V2X communications with significant improvements in reliability, data rate and wireless access [7].

Recent studies [7]–[10] manifest that 6G could bring substantial enhancements to V2X communications compared to its predecessors in the following aspects: (1) Ultra-high data rate is expected for faster data exchange in advanced applications like ultra-high-definition video streaming and augmented reality. Key technologies include millimeter-wave, visible light, and THz communications. (2) Ultra-low latency can reduce communication delay between V2X devices. This is crucial for time-critical applications, such as autonomous driving, where split-second decisions need to be made. Multiple medium access techniques, new multi-carrier schemes, and advanced resource allocation are the key supporting technologies. (3) Massive device connectivity is to maintain enormous connections among vehicles and other devices simultaneously. Technologies like non-orthogonal multiple access (NOMA), UAV communication, and satellite communication enable seamless and ubiquitous connectivity for V2X applications. (4) Advanced positioning and sensing can achieve centimeter-level positioning accuracy, which may be required by certain 6G V2X applications. (5) 6G could accommodate increasing demands for V2X communications while maintaining reliable connectivity and performance by utilizing IRS [5], [6] to enhance the spectral efficiency.

Evidently, many advancements in 6G V2X presented above rely on fine-grained, intelligent resource allocation for physical-layer resources. The resources in this study comprise of both time resource and frequency resource. To better understand it, we use a resource grid to denote them, as shown in Fig. 1. The resource grid has a two-dimensional structure at the physical layer. It consists of a control channel and a

Minglong Zhang and Vuk Marojevic are with the Department of Electrical and Computer Engineering, Mississippi State University, Mississippi, USA, e-mail: mz354@msstate.edu, vuk.marojevic@ece.msstate.edu.

Yi Dou is with the School of Computer Science, Nanjing University of Posts and Telecommunications, e-mail: yi.dou@njupt.edu.cn.

Peter Han Joo Chong is with the Department of Electrical and Electronic Engineering, Auckland University of Technology, Auckland, New Zealand, e-mail: peter.chong@aut.ac.nz.

Henry C. B. Chan is with the Department of Computing, The Hong Kong Polytechnic University, Hong Kong, e-mail: henry.chan.comp@polyu.edu.hk.

Authorized licensed use limited to: Mississippi State University Libraries. Downloaded on September 11,2023 at 23:37:11 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3294279

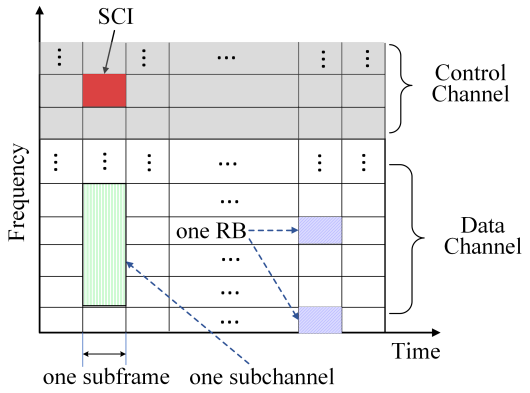


Fig. 1: A two-dimensional resource grid in 6G V2X.

data channel. Its basic unit is resource blocks (RBs). In the time domain, each RB is one subframe, which corresponds to a number of orthogonal frequency division multiplexing (OFDM) symbols. In the frequency domain, a number of subcarriers constitute a RB. Although one RB is the smallest unit that can be assigned to UEs, subchannels grouped by RBs are usually adopted in resource allocation. The number of RBs in a subchannel is configurable. Data from UEs are sent over data channel, while the control channel carries control information called sidelink control information (SCI). The SCI contains details for receiving data, such as which subchannels are used to carry the data in the data channel, modulation and coding scheme (MCS) used for the transmission, and other additional information.

In 5G V2X [2], two work modes (e.g., *Mode 1* and *Mode 2*) have been specified by the 3rd Generation Partnership Project (3GPP) to schedule sidelink transmissions and allocate physical-layer resources. *Mode 2* is an autonomous multiple access mode that vehicles can reserve physical-layer resources by themselves using a sensing-based semi-persistent scheduling (SPS) scheme. By contrast, in *Mode 1*, the resources are managed and allocated dynamically in a centralized way by gNodeBs.

Mode 1 facilitates satisfying massive communication demands if the allocation method is well designed, thereby ensuring the QoS and accommodating more transmissions at the same time. Unfortunately, few existing relevant research studies resource allocation in this mode. Instead, most of them focused on *Mode 2* [11]–[19]. These studies in the literature use either fixed rules, analytical models, or supervised learning. For instance, the authors [12] proposed an adaptive SPS scheme which allows each vehicle to flexibly adjust Resource Reservation Interval (RRI) according to the available channel resource. Machine learning has been widely leveraged to allocate resources for V2X communications [16]–[18], [20], [21]. Again, they are only applicable to the *Mode 2*. A few of studies have taken into account resource allocation for *Mode 1* [22]–[25], but their performance probably cannot meet the more stringent requirements in 6G V2X.

To satisfy the requirements, 6G V2X embraces advanced resource allocation schemes that are built to adjust allocations according to QoS feedback and context. Reinforcement learning (RL), which is a machine learning training method without the need for supervised learning and prior knowledge of the environment/system, has the potential to provide an intelligent resource allocation for 6G V2X. A RL system usually includes four elements: an agent/learner, an environment for the agent interacts with, a policy for governing the agent how to take actions and a reward given to the agent after taking an action. Through rewarding desired behaviours and/or punishing undesired ones, as well as sensing the state of the environment, a learning agent is able to discover the action policy that maximizes the cumulative reward to achieve an optimal decision/solution. RL algorithms can in general be classified into two types: model-free and model-based. While the former one is not based on an explicit model of Markov Decision Process (MDP), the latter one is based on a MDP model with clearly defined states, actions and rewards. Both are widely adopted in wireless communications, such as scheduling [26]–[28] and channel estimation [29], [30].

However, it is well known that there are exploration and exploitation issues for RL [7]. Hence, there is a need to develop fast-converging learning solutions for RL. This study strives for how to improve 6G V2X network performance by intelligently allocating resources while maintaining a fast convergence. As a consequence, a fuzzy-logic-assisted reinforcement learning model (named FAQ) is proposed to tackle the problems. The proposed FAQ algorithm intelligently and dynamically allocates physical-layer resources. It integrates the fuzzy logic and the Q learning to accelerate the learning. By leveraging the fuzzy logic, the FAQ allocates resources based on comprehensive evaluations while taking into account multiple metrics. It fully exploits the limited resources, thereby satisfying the stringent requirements of the V2X services orientated for low latency, ultra reliability and high data rate in 6G epoch. The FAQ allocation algorithm not only ensures the QoS for different V2X applications, but also has a faster convergence in learning, compared with other ML-based resource allocation strategies. Moreover, the algorithm also shows good scalability and applicability for vehicular networks with various vehicle densities and diverse V2X services.

The key contributions are as follows: 1) A novel resource allocation approach relying on reinforcement learning has been developed to maximize the network throughput through maximizing the reuse of the physical-layer resources; 2) Compared to other advanced learning, the FAQ not only has a faster convergence, but also outperforms other resource allocation schemes in throughput and packet delivery ratio; 3) An analytical model is developed to analyze the network throughput; 4) With the deployment of typical V2X services, a V2X simulation model in urban area is established to acquire the system performance and validate the mathematical model.

The rest of this paper is organized as follows. Section II reviews the related works for resource allocation in C-V2X. Section III introduces the system model and formulates the resource allocation problem. Section IV elaborates the proposed FAQ scheme. Section V provides theoretical analysis of the network performance for FAQ. Section VI evaluates the network performance and compares it with a few of benchmarks with advanced resource allocation. Finally, Section VII concludes the whole paper.

Authorized licensed use limited to: Mississippi State University Libraries. Downloaded on September 11,2023 at 23:37:11 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3294279

II. Related work

A. Centralized Resource Allocation for C-V2X

Since the 3GPP has no specific algorithm standardized for resource allocation in *Mode 1*, some existing scheduling schemes for C-V2X in this mode make use of geographical positions of a cluster of vehicles [31]–[33]. In [31], Abanto-Leon *et al.* propose a resource allocation scheme by grouping vehicles into clusters based on their locations. The authors indicate the transmission conflicts between vehicles related to the intra-cluster and inter-cluster interference. Paper [33] proposes a resource allocation scheme for road safety applications for V2X communications in *Mode 1*. Their scheme groups vehicles into different clusters. Each cluster is assigned an orthogonal resource set. Two clusters reuse the same resource set when the distance between their centroids is above a threshold. The mobility of vehicles challenge the periodicity of clustering and the stability of the cluster. Poor clustering algorithms would also influence the performance of resource allocation.

The other types of scheduling schemes in this mode consider the location of each vehicle individually and assign resources using the relative distance among vehicles [22]–[25], [34], [35]. Paper [22] proposes a location-aware resource allocation scheme (LARA) which allows limited number of concurrent links in the network while mitigating interference at recipients. The paper [24] showcases the benefits of network-orchestrated radio resource allocation mode with the proposed enhancements specifically targeting platooning applications. Multiple factors including locations of vehicles are processed using fuzzy logic to allocate resources in a comprehensive way in [25]. The study in [23] designs a context-based resource allocation scheduling scheme for LTE V2X *Mode 3* (similar to *Mode 1* in 5G V2X) which can reduce packet collision and half-duplex effects. By taking into consideration the geographical location of vehicles and dynamical configuration of the network operations based on the context condition, the scheme seeks to ensure that all vehicles experience similar interference level when the resources are shared. In fact, the majority of the resource allocation approaches in *Mode 3* are unable to guarantee the QoS, because a substantial portion of requests has been declined (especially in a dense network), which implies the corresponding UEs forfeit opportunities to exchange information with their peers. Regarding vehicle-to-vehicle (V2V) communications in the same mode, localization-based resource selection and scheduling scheme are presented in [34] and [35], respectively.

B. Decentralized Resource Allocation for C-V2X

Liang *et al.* [16] employ multi-agent reinforcement learning to address the spectrum sharing problem in vehicular networks, where multiple V2V links reuse the frequency spectrum preoccupied by vehicle-to-infrastructure (V2I) links. Yuan *et al.* develop a joint deep reinforcement learning (DRL)-based algorithm to enhance the performance of both V2I and V2V links [17]. To further provide the adaptiveness of resource allocation policy, the authors include meta-learning and develop a meta-based DRL for dynamic environments. In this work, the spectrum selection of V2V links uses *Mode 2*. Kim *et al.* provide a framework to measure the crash risk of a vehicle in the dynamically changing environment [18]. To optimize the operation of a vehicle adaptive to the environment, the authors presented a RL algorithm. The algorithm is designed to autonomously choose the optimal transport block size for the sidelink shared channel in *Mode 2*. The authors of [36] quantify four transmission errors and provide packet delivery ratio (PDR) against distance in LTE C-V2X *Mode 4* (similar to *Mode 2* in 5G V2X).

The authors of [20] investigated how to guarantee the QoS in cellular V2X. A joint transmission mode selection such as V2V or V2I communications, RB allocation and power control is considered. A deep reinforcement learning algorithm is proposed to maximize the capacity of V2I (i.e., vehicle-to-base station) links. However, its model adopts a shared resource pool between V2I and V2V links, which is evidently different with 5G V2X communications, where V2V and V2I have decoupled frequency bands. Wu *et al.* [37] designed a reinforcement learning-based data storage protocol to keep the information always in a specified region in vehicular Ad Hoc networks. The protocol adopts a fuzzy logic algorithm for a short-term assessment while choosing the next data carrier. However, the proposed fuzzy Q-learning is designed for maintaining the data in vehicular networks rather than improving the network performance in V2X communications. Vemiredd *et al.* [38] proposed a fuzzy reinforcement learning for the energy efficient task offloading in vehicular fog computing. The proposed scheduling algorithm combines reinforcement leaning and fuzzy logic based greedy heuristic. The algorithm aims to increase the offloading efficiency in vehicular networks rather than improving radio resource utilization.

C. Resource Allocation for 6G V2X

There exist a few of studies about radio resource allocation for V2X in the emerging 6G epoch. Paper [39] identifies that as one of the three technical pillars in cellular-based sidelink communications, resource allocation is prominent for reliable V2X communications. In the context of a large autonomous network, integrating space, air, ground, and underwater networks with ubiquitous and unlimited wireless connectivity, the authors in paper [40] point out that AI-enabled resource allocation can choose the most suitable scheduling for UEs at data-link layer for V2X and other types of communications. To alleviate insufficient resources and meet the demand for very high reliability and ultralow latency, the study propose an autonomous resource allocation (i.e., decentralized mode) enabled by NOMA for next-generation C-V2X [41]. It highlights the potential advantages gained from the NOMA-based scheme. In [42], a dynamic resource allocation method is presented for the base stations of the next-generation fully-decoupled C-V2X network. The method allocates appropriate bandwidth resources according to the various QoS requirements by applying the Lyapunov stochastic optimization method. Although the research [43] designs radio resource management empowered by federated Q-learning in 6G-V2X, it only considers how to choose one out of

Authorized licensed use limited to: Mississippi State University Libraries. Downloaded on September 11,2023 at 23:37:11 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3294279

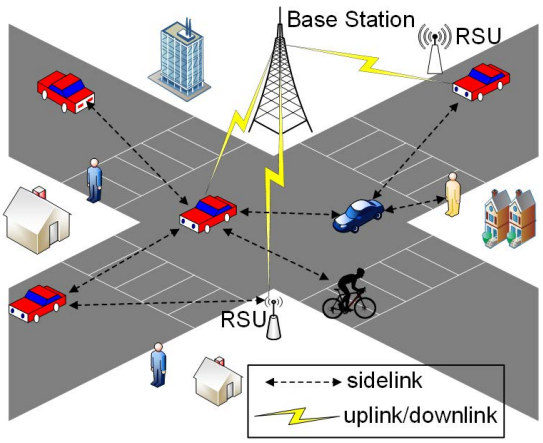


Fig. 2: Communication scenarios considered in the system model.

three alternatives, e.g., Dedicated Short Range Communication (DSRC), 5G millimeter-wave (mmWave) or 6G V2X, rather than assigning radio resources in a specific technique.

III. System Model and Problem Formulation

A. Network Model and Deployed Services

This study considers various V2X communication scenarios in a cellular network, as shown in Fig. 2. A centralized mode is considered, and the based station allocates physical resource blocks (PRBs) upon the reception of requests from UEs. The UEs here encompass both vehicular UE (VUE) and other smart devices, such as roadside unit (RSU) and cellular UEs (CUEs). The BS may decline a transmission request if it fails to find available PRBs within its permissible transmission window, i.e., exceeding the required latency for that particular V2X service. For each UE, two radio interfaces named *uU* and *PC5* are mounted on it. The former interface is a cellular interface to support vehicle-to-BS communications via the uplink/downlink, whereas the latter is for direct V2X communications via the sidelink. In addition, the interface *PC5* operates in a half-duplex mode for sidelink communications, which implies that UEs cannot send and receive data simultaneously via the interface.

We assume that six different V2X services are deployed in the system, as delineated in Table V in Appendix B. The table lists two basic V2X services, e.g., cooperative awareness message (CAM) and dynamic traffic control and warning. Defined by the European Telecommunications Standards Institution (ETSI) [44], CAM is a periodical safety message broadcast by all vehicles to maintain mutual awareness among them. Regarding the dynamic traffic control and warning, messages are periodically broadcast by RSUs to inform vehicles of the current road conditions and traffic situations. Two eV2X services, cooperative maneuver and cooperative sensing, are considered as well. The cooperative maneuver is to coordinate vehicles in some advanced driving cases, such as automatic platooning and assisted lane changing. Cooperative sensing refers to the extended sensors in eV2X. It is usually used to exchange data generated by multiple sensors mounted in vehicles when necessary or triggered by a certain event, to proactively prevent accidents. In this study, the transmission of cooperative sensing data is activated when vehicles are passing an intersection where collisions may take place due to complicated traffic situations. In fact, the two eV2X services require a much higher data rate than the basic ones. The last two use cases are both non-safety related services, which are conducted between RSUs and VUEs. The real-time content is widely used in social media and entertainment, such as audio-visual online chat and video streaming, while the non-real-time case is demanded by data downloading and uploading, such as sending/receiving emails, and advertisement delivery. For the convenience, we use SER1, SER2, ... , SER6 to refer to six services in the following discussions, as noted in the Table V in Appendix B.

TABLE I: Variable notation in the system model

Variable	Definition
PL	Total path loss of a wireless link
d, d_0	Distance between transmitter and receiver, Reference distance for measure
γ, X_σ	Path loss exponent, Random shadowing effects
P_t, P_r	Transmission power, Reception power
G_t, G_r	Antenna gain for transmission, reception
z, Ω	Received signal level, Average received signal level
m	Fading depth in Nakagami fading
r, z_r	Radius of UE's transmission range, Radius of a positioning zone
Z, PZ_i	The number of positioning zones, A positioning zone
$C, y_{i,j}$	Conflicting matrix, Conflicting degree
$X, x_{i,j}$	Allocation matrix, Allocation element
R	The maximum number of subchannels
TH^H, TH^M, TH^L	Throughput of high, medium, and low-priority services, respectively
$\Upsilon^H, \Upsilon^M, \Upsilon^L$	Successful allocation ratio of high, medium, and low-priority services, respectively

In the system, frequency division multiplexing access (FDMA) is used. In this multiplexing mode, UEs can be assigned subchannels at a particular subframe to access data channel. How to allocate subchannels and subframe for all transmission requests determines the network performance (e.g., network throughput and spectral efficiency), especially when the number of the transmission requests increases, resulting in more competitive situations for accessing the medium. The notations in the system model are listed in Table I.

From the 3GPP V2X specifications [45], the adopted link pathloss in this study is:

$$PL(d) = PL(d_0) + 10\gamma \log(d/d_0) + X_\sigma, \tag{1}$$

where PL is the total path loss measured in decibel (dB), $PL(d_0)$ is the path loss at the reference distance d_0, d is the distance between transmitter and receiver, γ is the path loss exponent and X_σ describes the random shadowing effects (a zero-mean, normally distributed random variable with standard deviation σ). Correspondingly, the received power P_r can be denoted by:

$$P_r = P_t + G_t + G_r - PL(d), \tag{2}$$

where P_t is the transmit power and G_t and G_r are the antenna gains in dBi. Nakagami-m distribution is suitable for

Authorized licensed use limited to: Mississippi State University Libraries. Downloaded on September 11,2023 at 23:37:11 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3294279

describing statistics of mobile radio transmission in complex medium such as the urban environment. We assume that all the V2X links in the system experience Nakagami fading, so the received signal level z follows:

$$f_z(z,\Omega) = \frac{2}{\Gamma(m)}(\frac{m}{\Omega})^m z^{2m-1}\exp(-\frac{mz^2}{\Omega}), z>0, m \geq \frac{1}{2}, \tag{3}$$

where $\Gamma(\cdot)$ is the gamma function, $\Omega = E[z^2]$ is average signal power, and m is the fading depth defined by

$$m = \frac{E^2[z]}{Var[z^2]}. \tag{4}$$

B. Distribution Pattern and Conflicting Zones

When considering how the subchannels are reused, the geographical locations among vehicles, their distances and request types are factored in. Therefore, we use a distribution pattern to model both the dynamics of vehicles and their requests. First, the service area covered by the base station is partitioned into multiple positioning zones, as shown in Fig. 3. Then the pattern is defined as a tuples consisting of a positioning zone's location and the requests sent by the vehicles in the zone. The purpose of partitioning into zones is to represent the location of vehicles by the positioning zones they belong to, instead of their accurate geographical locations. Owning to the movement of vehicles, the requests for service also keep changing, which further brings forth the changes of distribution patterns.

We assume that the radius of each positioning zone is z_r, and that the radius of coverage area for each positioning zone is r ($z_r << r$), which is also transmission range of CUEs and VUEs. The definition of conflicting zones is that their coverage areas are overlapped. That is, any two non-conflicting zones are at least r/z_r zones apart. All vehicles' requests issued from the same conflicting zones will interfere with each other if the same resources are allocated. On the other hand, when the requests are issued from non-conflicting zones, they can reuse the same subchannel and subframe. Fig. 4 shows the relationship of the positioning zones, the transmission range and the conflicting zones in a general Manhattan grid.

Assume that there are M vehicles and Z positioning zones. Some vehicles may involve in several different conflicting zones. We define a conflicting matrix C of all vehicles as follows:

$$C = \begin{bmatrix} y_{1,1} & y_{1,2} & \cdots & y_{1,Z} \\ y_{2,1} & y_{2,2} & \cdots & y_{2,Z} \\ \cdots & & \cdots & \\ y_{M,1} & y_{M,2} & \cdots & y_{M,Z} \end{bmatrix}. \tag{5}$$

Matrix C shows conflict relations between requests issued by vehicles in the different zones. $C_{i,*} = [y_{i,1}, ..., y_{i,Z}]$ denotes the i-th row of the matrix C representing the conflicting zones for vehicle v_i. A generalized format of the conflicting degree is defined according to the distance between two zones. The following equation shows a normalized result, in which all values of the degree are real numbers between [0, 1]:

$$y_{i,j} = \begin{cases} 0, \ d_{ij} > 2r \\ 1 - \dfrac{d_{ij}}{2r}, \ d_{ij} \in [0, 2r] \end{cases} \tag{6}$$

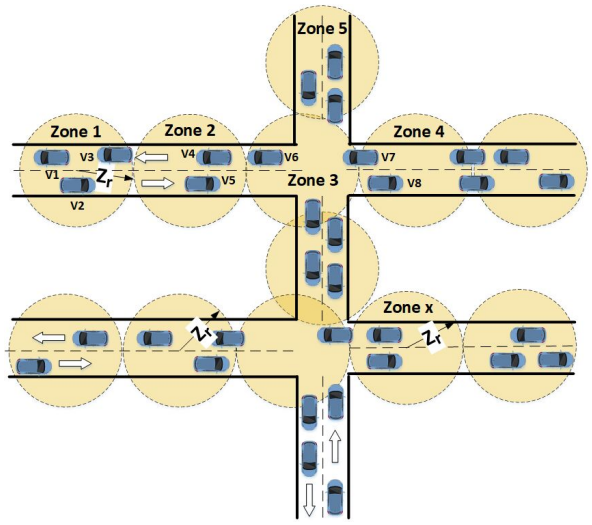


Fig. 3: Partitioning of the entire area into multiple positioning zones.

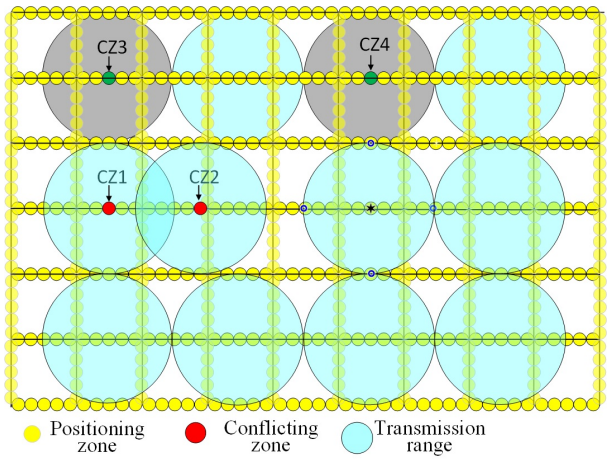


Fig. 4: Demonstration of conflicting zones in a general Manhattan grid: zone CZ1 and CZ2 in red are conflicting zones, while the two gray zones CZ3 and CZ4 are non-conflicting ones.

where d_{ij} is the distance between vehicle v_i and the center of a positioning zone PZ_j, and r is the radius of the transmission range. Note that the minimum distance between two vehicles is zero. If $y_{i,j}$ is zero, it implies that a vehicle v_i does not conflict with vehicles in PZ_j. Otherwise, they may have conflicts. The shorter distance between the vehicle v_i and the zone PZ_j is, the larger degree of conflicting will be caused. Without losing generality, the authors select a binary format in this study. If the distance $d_{ij} \geq 2r$, $y_{i,j} = 0$. Otherwise, $y_{i,j} = 1$.

Except for general road networks like Manhattan grid, the proposed FAQ is applicable to irregular road patterns because the partitioning method in Fig. 3 and conflicting zone definition in Fig. 4 can be directly applied in an irregular road network. After determining conflicting zones, the FAQ algorithm is able to formulate a state space and allocate resources according to their conflicting relationships.

Authorized licensed use limited to: Mississippi State University Libraries. Downloaded on September 11,2023 at 23:37:11 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3294279

C. Problem Formulation

The aim is to maximize the network throughput by properly assigning PRBs to various UEs in the network. Assume there are M UEs, and each UE can obtain R subchannels at most. We first define a resource allocation matrix X:

$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & ... & x_{1,R} \\ x_{2,1} & x_{2,2} & ... & x_{2,R} \\ ... & & ... & \\ x_{M,1} & x_{M,2} & ... & x_{M,R} \end{bmatrix}. \quad (7)$$

Each $x_{i,j}$ indicates whether the j-th subchannel ($1 \leq j \leq R$) is allocated to the i-th UE ($1 \leq i \leq M$) as shown below:

$$x_{i,j} = \begin{cases} 1, \text{ if subchannel } j \text{ is allocated to UE } u_i \\ 0, \text{ otherwise} \end{cases}. \quad (8)$$

Let TH^H, TH^M, and TH^L represent the throughput of high, medium, and low-priority services[1] in the network, respectively. The resource allocation optimization problem can be formulated as follows:

$$X^{opt} = arg \max_{x_{i,j} \in X} TH^H + arg \max_{x_{i,j} \in X} TH^M + arg \max_{x_{i,j} \in X} TH^L \quad (9)$$

subject to

$$\Upsilon^H \geq \Upsilon^M \geq \Upsilon^L, \quad (10)$$

$$\sum_j x_{i,j} = k_i,$$

$$x_{i,j} = x_{i,j+1} = ... = x_{i,j+k_i-1} = 1 \ (0 \leq k_i \leq R), \quad (11)$$

$$\text{and } T^{sch}_{u_i} \leq p^{tx}_{u_i} + t^{req}_{u_i}. \quad (12)$$

The term X^{opt} in (9) is the optimal allocation. The equation (10) shows the constraint that the successful allocation ratios (i.e., $\Upsilon^H, \Upsilon^M, \Upsilon^L$) should descend from high-priority services to low-priority services, which implies services with higher priority have advantage to get resources against the ones with low priority. Equation (11) is set to guarantee that k_i consecutive subchannels are assigned to UE u_i. Equation (12) ensures that the scheduled transmission moment $T^{sch}_{u_i}$ must be no later than the allowable latency, that means it is less than the sum of transmission period $p^{tx}_{u_i}$ and the request arrival time $t^{req}_{u_i}$.

To reach the maximum network throughput, optimal allocation is required. Note the matrix X in (7), if each UE requires k subchannels ($k << R$), there will be $(R-k)^M$ possible allocations. Given a vehicular network having hundreds or even thousands of vehicles and $R = 20$ subchannels in the resource pool, the possibilities are nearly infinite, which implies that we cannot acquire an analytic form for the optimal X. Therefore, we propose an advanced reinforcement learning model to find the sub-optimal allocations.

[1]According to their properties, the considered six V2X services in this system have different levels of priority, which will be elaborated in Table II in section IV.

IV. FAQ Model for Resource Allocation

A. Overall Architecture of the Proposed FAQ Algorithm

The overall architecture of the proposed fuzzy-logic-assisted learning model is illustrated in Fig. 5. The ultimate goal of the learning model is to figure out the best allocation for each vehicle distribution pattern. In this study, a Q-learning model is embedded and trained to learn how to allocate resources.

The Q-learning model repeatedly adjusts the allocations for all requests in the buffer until obtain an optimal resource allocation strategy. It involves a state set S and an action set A for each state to describe the interactions between an agent and an environment. The environment is the base station and the communication network among UEs. The state represents the features or situations about the network and the available resources. The set of actions are the options that the base station can choose to modify the allocations. When the base station performs an action $a_t \in A$ in state $s_t \in S$, it obtains a new network state s_{t+1} and a reward $r(s_t, a_t)$ as the "feedback" of the V2X communications.

The *action selector* selects actions (i.e., allocate resources by making a decision) for transmission requests using one of the two principles: either fuzzy logic strategy or ϵ-greedy algorithm. The fuzzy logic is able to accelerate the learning and convergence by taking into account multiple metrics (e.g., timing, interference, mutual reception, priority and conflict zones), whilst the latter explores more options and considers the long-term learning outcomes. The *reward calculation* collects data and calculates the corresponding reward (i.e., network throughput) induced by executing the elected actions.

To evaluate the potential long-term reward, the FAQ model formulates $Q(s_t, a_t)$ for each state-action combination which is computed as the expected total rewards of all future adjustments starting from the current situation. In the learning phase, the base station adopts a trial and error training procedure to update $Q(s_t, a_t)$ according to the reward attained. The *action policy adjustment* is responsible for dynamically tuning parameters that determine which strategy the *action selector* chooses to allocate resources. Once the updating of $Q(s_t, a_t)$ has converged, the agent chooses the resource allocation that provides the highest expected future reward.

B. Fuzzy-Logic Based Action Selection

Fuzzy logic is one of the approaches leveraged for action selection during the learning process. It is implemented in the module *Fuzzy logic process*. The module considers four factors comprehensively to judge whether a new tentative allocation is better or worse. A quantitative result will be output at the final stage for decision making. If the allocation is better than the current one, it will be adopted. Otherwise, another allocation will be assessed. The considered factors are timing, interference, mutual reception ratio, and service priority. The procedure is as follows:

1) Fuzzification: We use predefined linguistic variables and membership functions to convert the above factors to fuzzy values ranging from 0 to 1.

- Timing factor
  The timing factor denotes the extent of urgency for a

Authorized licensed use limited to: Mississippi State University Libraries. Downloaded on September 11,2023 at 23:37:11 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3294279

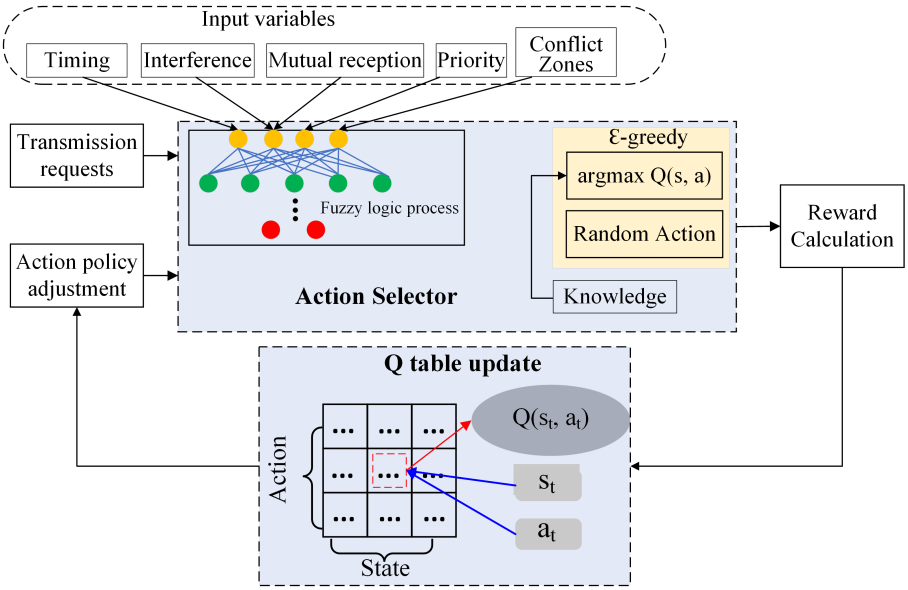


Fig. 5: Architecture of the proposed FAQ learning model.

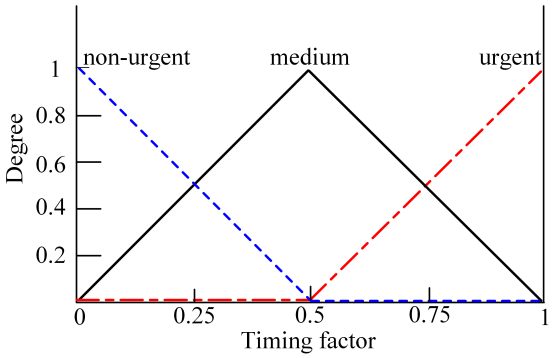


Fig. 6: Timing factor membership function

transmission, and guarantees that the scheduled time for a transmission request will not exceed the end-to-end latency. Assume that the largest permissible delay for a service is P, the timing factor is defined as:

$$T(w) = \frac{w}{P}, \ (1 \leq w \leq P), \tag{13}$$

where w the sequential number of the planning subframe in the transmission window. The higher the value, more urgent a transmission becomes. Fig. 6 illustrates its membership function.

- Interference factor
  Regarding a node n with a transmission request, the local signal-to-interference-plus-noise ratio (SINR) at the recipient is

$$SINR_n = \frac{P_{rx}^n}{P_N + \sum_{i=1}^{N_{inf}} P_i^{inf}}, \tag{14}$$

where P_{rx}^n is the received power at the recipient, P_N is the noise power and P_i^{inf} is the interference power from node i. They can be acquired according to (1) to (3). In fact, interference is reciprocal. When other VUEs pose interference to a requesting VUE, the interfered VUE also inversely cause interference to all other VUEs at the same time. Thus, for each VUE, we can calculate the local SINR correspondingly. To evaluate the global interference level for the current circumstance and assume there are K VUEs interfering each other, the average interference level will be the overall interference divided by K:

$$SINR_{avg} = \frac{\sum_{i=1}^{K} SINR_i}{K}. \tag{15}$$

The interference factor membership function is shown in Fig. 7, and the corresponding expressions for high-interference, medium-interference and low-interference curves are:

– the low interference:

$$f_{IL}(x) = \begin{cases} 1, & x \geq f_c \\ \frac{x-f_b}{f_c-f_b}, & x \in [f_b, f_c) \\ 0, & x < f_b \end{cases} \tag{16a}$$

– the medium interference:

$$f_{IM}(x) = \begin{cases} \frac{x-f_a}{f_b-f_a}, & x \in [f_a, f_b) \\ \frac{f_c-x}{f_c-f_b}, & x \in [f_b, f_c) \\ 0, & other \end{cases} \tag{16b}$$

– the high interference:

$$f_{IH}(x) = \begin{cases} 1, & x < f_a \\ \frac{f_b-x}{f_b-f_a}, & x \in [f_a, f_b) \\ 0, & x \geq f_b \end{cases} \tag{16c}$$

- Mutual reception factor
  Some V2X services, such as CAM, require UEs receive each other's messages at the same time. For instance, if two VUEs are within each other's communication range and broadcast CAM messages concurrently via different subchannels, vehicles around them can receive the information from both, but the two sending nodes cannot receive each other's information due to the attribute of the half-duplex radio. Such a situation diverges

Authorized licensed use limited to: Mississippi State University Libraries. Downloaded on September 11,2023 at 23:37:11 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3294279

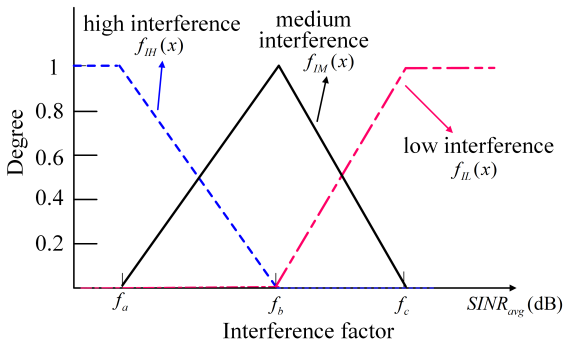


Fig. 7: Interference factor membership function.

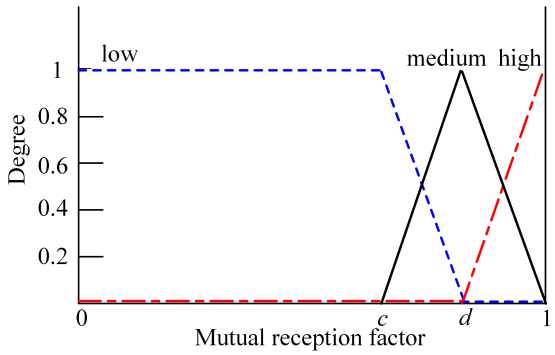


Fig. 8: Mutual reception factor membership function.

from the necessity of certain applications, such as mutual awareness delivered by CAM messages, and should be avoided. On the other hand, if two transmitters lie outside each other's transmission range and they use different subchannels to broadcast simultaneously, other vehicles in the overlapped area can receive the CAM messages from both vehicles, which in turn enhances the road safety. Suppose r is the transmission range of a UE, for a request ς [2] from node n and a request ρ from node m ($m \neq n$), the definition of mutual reception factor is

$$M_{\varsigma,\rho}(n,m) = \begin{cases} 1, & d(m,n) \leq r \\ \frac{r^{\gamma}}{d(m,n)^{\gamma}}, & d(m,n) > r \end{cases} \quad (17)$$

where γ is the path loss component in the channel model. The average value of mutual reception factor for request ς at node n is

$$M_{\varsigma}^{avg} = \frac{\sum_{\rho=1}^{H} M_{\varsigma,\rho}(n,m)}{H}, \quad (18)$$

where H represents all the requests excluding the request ς. The average level of the mutual reception factor M_{ς}^{avg} will be the input of the membership function, as shown in Fig. 8.

- Priority factor
  Different V2X services have different priority levels. According to [1], safety-related services have higher priority than non-safety counterparts, and hence we prioritize the eV2X services at the highest level, and non-safety services at the lowest level. In our system model, there is a total of six types of services coexisting in the network. The priority factor and channel bandwidth BW (i.e., required number of successive subchannels) are summarized in Table II.

TABLE II: Priority factor

Service Category	Priority	Number of subchannels
Cooperative manoeuvre (SER2)	High	10
Enhanced sensing (SER3)		18
Cooperative awareness (SER1)	Medium	3
Dynamic traffic control (SER4)		8
Non-safety non-realtime service (SER5)	Low	12
Non-safety realtime service (SER6)		

[2]A node may have more than one requests with different types.

2) Inference: Apply the predefined IF/THEN rules in Table III to reach a verdict and rank the combinations of rules according to a min-max principle detailed in [25].

TABLE III: Applied Fuzzy Rules

Rule	Timing	Interference /Mutual Reception	Priority	Verdict
1			Low	Not Acceptable
2	Non-urgent	Low	Medium	Acceptable
3			High	Good
4			Low	Bad
5	Non-urgent	Medium	Medium	Not Acceptable
6			High	Acceptable
7			Low	Bad
8	Non-urgent	High	Medium	Bad
9			High	Not Acceptable
10			Low	Acceptable
11	Medium	Low	Medium	Good
12			High	Perfect
13			Low	Not Acceptable
14	Medium	Medium	Medium	Acceptable
15			High	Acceptable
16			Low	Bad
17	Medium	High	Medium	Bad
18			High	Not Acceptable
19			Low	Acceptable
20	Urgent	Low	Medium	Good
21			High	Perfect
22			Low	Not Acceptable
23	Urgent	Medium	Medium	Acceptable
24			High	Good
25			Low	Bad
26	Urgent	High	Medium	Bad
27			High	Not Acceptable

3) Defuzzification: Fuzzy values from the module *Inference* will be converted to numerical values during the process of defuzzification, by applying an output membership function illustrated in Fig. 9 and a defuzzification method named center of gravity (COG) [25]. The numerical output implies applicability of the current RBs for a request. A higher value means more appropriate. A threshold $defuz$ is set for determining whether the resources will be allocated to the requesting UE.

C. Fuzzy-logic-Assisted Q Learning Process

1) Initial resource allocation using fuzzy logic: We first use the fuzzy logic approach proposed in Section IV-B to obtain an initial resource allocation for all requests that sent by different UEs in the same allocation session, then we employ the Q learning to improve the initial allocations.

Authorized licensed use limited to: Mississippi State University Libraries. Downloaded on September 11,2023 at 23:37:11 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3294279

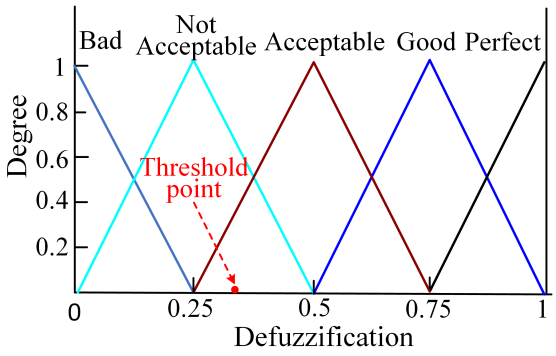


Fig. 9: Output membership function for resource selection

2) State: We formulate state s_t as a vector describing the existing resource allocations or available resources at the base station for the t^{th} transmission request stored in the buffer of the base station. That is

$$s_t = \{X^{t-1}, C_{t,*}, (\Upsilon^H, \Upsilon^M, \Upsilon^L), (TH^H, TH^M, TH^L)\},$$

where X^{t-1} represents resource allocation of all requests after $t-1$ times of adjustments. X^{t-1} summarizes the entire sequences of all past adjustments that led to it. X^0 denotes the initial resource allocation of all requests using the fuzzy logic approach. $C_{t,*}$ represents the conflicting zones of vehicle v_t issuing the t^{th} request defined in Equation (5). By applying resource allocation X^{t-1}, $(\Upsilon^H, \Upsilon^M, \Upsilon^L)$ means the successful allocation ratios of all requests and (TH^H, TH^M, TH^L) stands for the network throughput obtained after all requested communications have been completed.

State s_t matters the future allocations adjustments. It does not matter how the allocations in s_t came about. Thus, the state transitions in our model have the Markov property that can predict the next state and expected reward based on the knowledge of the current state and action. Fig. 10 shows an example of state transitions. We adjust request req_1 in the subframe 1 to the subchannel ranges [4 $\sim$ 6]. The request req_2 is adjusted from unallocated to assigned with subchannel ranges [1 $\sim$ 8] in subframe 2. The request req_3 in subframe 3 is moved to subchannel ranges [7 $\sim$ 9].

3) Action: We formulate the action as a tuple $a_t = \{SER_{req}, (sch, sfr)\}$. The first element indicates the type of request being adjusted. The second element is a pair of subchannel sch and subframe sfr that are selected by the base station for the request, which is a new selection of resource blocks after the adjustment. The action indicates that the base station adjusts the subchannel and subframe allocation of a request. The action space includes all possible pairs of subchannels and subframes assigned to the requests. In our model, there are 7 possible successive subchannels choices at most for each of the 20 subframes. The action space size is 140 allocations of resource blocks (i.e., 7 * 20 = 140). As illustrated in Fig. 16, different types of requests have different choices of successive subchannels. For instance, the type of SER2 requests have at most 5 possible successive subchannels choices. Thus, for SER2 type requests have a total 100 allocation choices in the actions space. In the example of Fig. 10, req_1 with type of SER1 is adjusted to subchannel 4 in subframe 1. Thus, its action is written as $a_1 = \{SER_1, (sch: 4,\ sfr: 1)\}$. The actions of requests req_2 and req_3 are $a_2 = \{SER_4, (sch: 1,\ sfr: 2)\}$ and $a_3 = \{SER_1, (sch: 7,\ sfr: 3)\}$, respectively.

Algorithm 1: The FAQ Learning Algorithm

```
Initialize all Q(s,a) values;
Initialize θ, γ and ϵ;
for each episode do
    Initialize subchannel and subframe allocations for
     all requests based on the Fuzzy logic algorithm;
    Initialize starting state;
    for each request do
        if random (0,1)< ϵ then
            choose an action a_t for state s_t depending
             on the assessment results from the fuzzy
             logic;
        else if ϵ <random (0,1)< η then
            choose an action yielding highest Q-value
             for state s_t: a_t = argmax_{a∈A} Q(s_t, a) ;
        else
            choose an action a_t for state s_t randomly;
        end
        Execute the selected action a_t, observe instant
         reward r(s_t, a_t) and next state s_{t+1};
        Update Q(s_t, a_t) value using Equation (20);
        Current state s_t ← s_{t+1};
    end
end
```

4) Instant Reward: The instant reward $r(s_t, a_t)$ is the network performance (i.e., network throughput) of all requests in the buffer based on each allocation. As defined in (9), (10), (11), and (12), the resource allocation aims at maximizing the network throughput under the constraint of correct allocation ratios of different priority service requests:

$$r(s_t, a_t) = TH^H + TH^M + TH^L. \tag{19}$$

Here, the reward evaluates whether each request allocation adjustment increases or reduces the entire throughput of all requests. Unlike the traditional Q learning model, our instant reward is attained after a communication delay rather than in an immediate way. The performance result is obtained when the communications of all requests have been met.

5) Update the Q-table and the learning process: In the learning phase, the base station gradually adjusts the resource allocations to achieve a better network performance. During this process, the Q-value Q_t is used as output for making the decisions. In the state t, the base station selects an action a_t and gets an instant reward $r(s_t, a_t)$. The Q table is updated for this state action pair as

$$Q(s_t, a_t) = (1-\theta)Q(s_t, a_t) + \theta\{r(s_t, a_t) + \gamma \max_a Q(s_{t+1}, a)\}. \tag{20}$$

In the above, θ and γ are the learning rate and the discount rate, respectively.

Authorized licensed use limited to: Mississippi State University Libraries. Downloaded on September 11,2023 at 23:37:11 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3294279

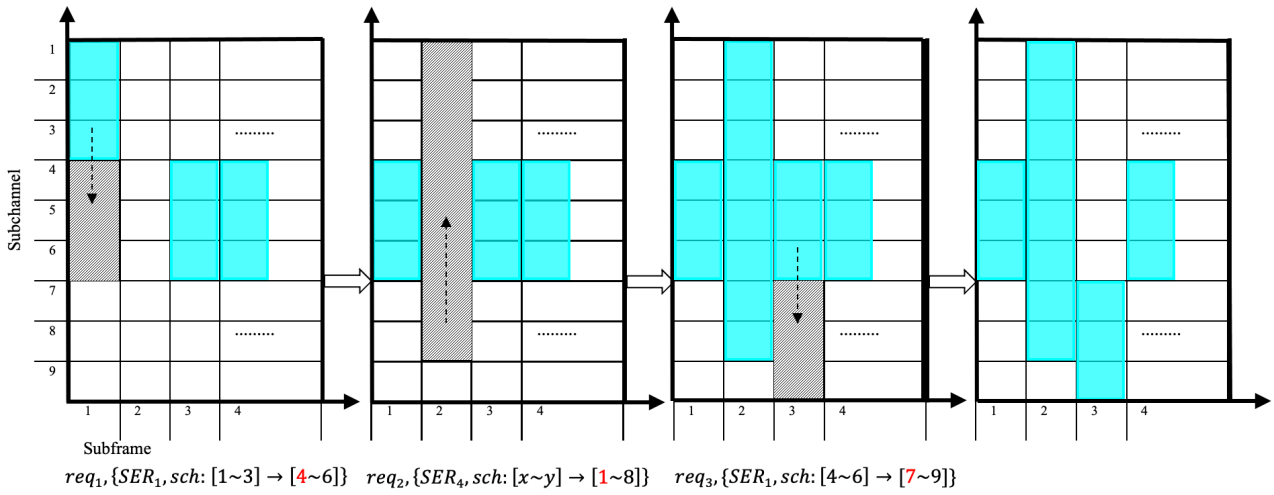


Fig. 10: State transitions by adjusting the allocations.

6) Action policy adjustment: Any action chosen by the base station is within the action space. However, the conventional strategy is the ϵ greedy so as to cover all possible actions. That is, with ϵ probability, the base station randomly selects an action from the action space, otherwise with $1-\epsilon$ probability, the agent chooses the action that has the highest Q-value based on the previous learning iterations. This strategy allows the base station to explore all possible selections until finding the best one. As opposed to the conventional greedy-based strategy, in the FAQ model, the agent will not randomly choose one action from the corresponding action spaces. Instead, the fuzzy logic is leveraged to intervene in the action selection, as shown in lines 7-13 of Algorithm 1.

At each iteration of learning, instead of blindly choosing one action from action space A, the agent applies fuzzy logic to grade potential allocations and the results, the defuzzification outputs, are compared. An inferior allocation will be replaced by a better one, if the agent can find a more appropriate allocation with higher defuzzification value. If not, it will choose the existing allocation which means no adjustment is made in the current state. Such mechanism can rule out some unreasonable decisions, thereby accelerating the learning, because the fuzzy logic based action selector utilizes multiple criteria to make a comprehensive evaluation for each tentative allocation.

The probability to choose an action is dynamically adjusted at different learning stages based on the observation of network throughput, rather than being kept unchanged parameters. ϵ and η are initialized as 0.5 and 0.8, respectively. With the learning in process, ϵ and η gradually increase but keeping $\epsilon < \eta$ if a throughput increase is observed. This makes the throughput level off at a higher value without further fluctuations.

V. Theoretical Analysis for Network Throughput

In this section, we conduct the theoretical analysis for the network throughput, based on the communication models in Section III and inspired by [46]. We first derive the bit error rate (BER) and then the packet error rate (PER). By considering resource allocation matrix X (see (7)), the network throughput can be finally obtained.

A. BER Analysis for V2X

At the PHY layer, SC-FDMA is adopted for sidelink communications. A transmitter precodes complex symbols (BPSK or M-QAM) by means of a discrete Fourier transform (DFT) operation before mapping them onto N_c allocated subcarriers by means of the Inverse Fast Fourier transform (IFFT). The mapping matrix L is defined as $L_{i,j} = 1$ if the symbol j is transmitted over subcarrier i and zero otherwise. At the reception side, a receiver first discards the cyclic prefix (CP) that has been added before transmission, and then performs the FFT that converts each time-domain symbol into a frequency-domain symbol before applying the $M-ary$ DFT that yields Y_M. After performing for demapping by L^H, the N allocated subcarriers Y_N are given

$$Y_N = L^H H_M L F S + L^H \eta_M, \tag{21}$$

where S is the transmitted symbols, the Fourier matrix F is defined as $F_{j,k} = \exp(2\pi \jmath/N_c)jk/\sqrt{N_c}$, η_M is the noise vector whose elements are random variables characterized by i.i.d. complex normal distribution $CN(0, N_0)$. Correspondingly, a $M \times M$ diagonal matrix H_M can be used to represent the channel frequency response for each subcarrier. Note that the matrix elements are complex circularly symmetric random variables.

The expression for the recovered symbol after zero-forcing frequency-domain equalization (ZF-FDE) yields

$$\tilde{s} = W F^H F S + W F^H \eta, \tag{22}$$

where W is an $N_c \times N_c$ complex matrix $W = F(H^H H)^{-1} H^H F$. The expression for the kth received symbol after ZF-FDE is

$$\tilde{s_k} = s_k + \Sigma_{j=1}^{N} \hat{\eta}_{j,k} = s_k + \tilde{\eta}_k, \tag{23}$$

where $\hat{\eta}_{j,k} = \frac{F^*_{j,k}\eta_j}{h_j}$.

It is observed that, after ZF-FDE, a received symbol is generated by adding the transmitted symbol to an effective noise term $\tilde{\eta}_k$. That means, the effective noise is the result of adding an enhanced Gaussian noise $\hat{\eta}_{j,k}$ to each assigned subcarrier.

Authorized licensed use limited to: Mississippi State University Libraries. Downloaded on September 11,2023 at 23:37:11 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3294279

As the effective SNR is conditional to the channel frequency response, we have $SNR = \frac{E_s h^2}{N_0}$, where N_0 and E_s are the noise power spectral density and the transmitted energy per symbol, respectively.

The term η_j is complex Gaussian noise, so the length of the two-dimensional vector follows a Rayleigh distribution. After multiplying η_j with $F^*_{j,k}$, the joint probability density function (PDF) is

$$f(x) = \frac{1}{2\pi}\frac{x}{\sigma^2}e^{\frac{-x^2}{2\sigma^2}} \tag{24}$$

with variance $\sigma^2 = \frac{N_0}{\sqrt{Nc}}$.

Since each effective noise element is a circularly symmetric random variable, the sum also obeys the property of circular symmetry, and its marginal distributions share the same even function. Taking into consideration this property, we assume the independence among elementary noise terms and derive the PDF $\tilde{\eta}_k$ as

$$p_{\tilde{\eta}_k}(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\Psi_{\tilde{\eta}_k}(\omega)e^{j\omega x}d\omega. \tag{25}$$

Considering $\sigma^2 = \frac{N_0}{\sqrt{Nc}}$ and $SNR = \frac{E_s h^2}{N_0}$, we get $\sigma^2 = \frac{E_s h^2}{\sqrt{Nc}SNR}$. Thus,

$$\begin{aligned}\Psi_{\tilde{\eta}_k}(\omega) = &\left(\frac{\left(\frac{E_s h^2}{SNR\sqrt{Nc}}\right)^{m/2}|\omega|^m}{2^{m/2-1}\Gamma(m)}\right)\left(\frac{m}{\Omega}\right)^{m/2} \\ &K_m\left(\left(\frac{E_s h^2}{\sqrt{Nc}SNR}\right)^{1/2}\sqrt{\frac{2m}{\Omega}|\omega|}\right)^N.\end{aligned} \tag{26}$$

For square M-QAM encoded modulation with Gray mapping [46], note that the effective noise term $\tilde{\eta}_r$ is circularly symmetric, so the BER can be calculated by

$$BER = \sum_{n=1}^{L-1} w(n)I(n), \tag{27}$$

where $w(n)$ is the coefficient dependent on the constellation mapping (i.e., $L = 2$ for BPSK, and $L = \log_2(M)$ for M-QAM). According to [47], for Gray mapping and square QAM,

$$\begin{aligned}w(n) = &\frac{1}{M\log_2(M)}\left[\sum_{u=1}^{\sqrt{M}-n}\alpha_u^+(u+n-1)\right. \\ &\left.+\alpha_{u+n}^-(u)+\beta_u^+(u+n-1)+\beta_{u+n}^-(u)\right],\end{aligned} \tag{28}$$

where

$$\alpha_u^{\pm}(k) = \pm\frac{1}{2}\Sigma_{i=1}^{\log_2(\sqrt{M})}\Omega(i,u)[\Omega(i,k)-\Omega(i,-k)] \tag{29}$$

and

$$\Omega(i,x) \triangleq sign\left[\cos\left(\frac{\pi}{2^i}\left(x-\frac{1}{2}\right)\right)\right]. \tag{30}$$

Denote $I(n)$ as the components of error probability (CEPs), which can be expressed as

$$I(n) = P_r\{\Re(\tilde{\eta}) > (2n-1)d\} = 1 - F_{\tilde{\eta}}((2n-1)d), \tag{31}$$

where d is the minimum distance between each modulated symbol and the decision boundary. According to the constellation, for BPSK and M-QAM, we have $d = \sqrt{E_s}$ and $d = \sqrt{(3E_s/2(M-1))}$, respectively.

The cumulative distribution function (CDF) can be calculated as [46]

$$\begin{aligned}F_{\tilde{\eta}} \approx &\frac{1}{2} + \frac{1}{2\pi}\left(\frac{\omega_{max}-\omega_{min}}{n}\left[\frac{\chi(x,\omega_{min})+\chi(x,\omega_{max})}{2}\right.\right. \\ &\left.\left.+\sum_{k=1}^{n-1}\chi\left(x,\omega_{min}+k\frac{\omega_{max}-\omega_{min}}{n}\right)\right]\right),\end{aligned} \tag{32}$$

where the auxiliary function $\chi(x,\omega)$ is defined as

$$\chi(x,\omega) = \frac{e^{jx\omega}\Psi_{\tilde{\eta}_k}(-\omega) - e^{-jx\omega}\Psi_{\tilde{\eta}_k}(\omega)}{j\omega}. \tag{33}$$

In (32), w_{min} is the minimum value in the integration interval, and it is set to a positive number very close to zero (e.g., 10^{-15}). We elect a value for w_{max} so that $w_{max} = \underset{w}{arg\,max}(\Psi_{\tilde{\eta}_k}(\omega) \leq 10^{-9})$. Finally, by substituting equations from (28) to (32) into (27), we obtain an approximate closed-form expression for the BER

$$\begin{aligned}BER \approx &\sum_{n-1}^{L-1} w(n)\left(\frac{1}{2} - \frac{\omega_{max}}{2\pi n}\left(\frac{1+\chi((2n-1)d,\omega_{max})}{2}\right.\right. \\ &\left.\left.+\sum_{k=1}^{n-1}\chi\left((2n-1)d,k\frac{\omega_{max}}{n}\right)\right)\right).\end{aligned} \tag{34}$$

The analytical results for BER performance are shown in Fig. 11 (a). The BER values for QPSK, 16-QAM and 64-QAM are acquired at different SNRs, with the number of subcarriers $N_c = 64$. The BER values will be utilized to calculate the packet error rate and network throughput.

B. Analysis of Network Throughput

To calculate the interference at a recipient vehicle v_j, we define a channel gain vector by $G_{1j} = \begin{pmatrix} g_{1j} & g_{2j} & g_{3j} & \cdots & g_{Mj}\end{pmatrix}$, where g_{xj} $(1 \leq x \leq M)$ denotes the channel gain between a sender v_x and a receiver v_j. Element g_{jj} is zero because when v_j is transmitting, it cannot receive itself at the same time due to the half-duplex property. Thus, we define a $M \times M$ channel gain matrix as:

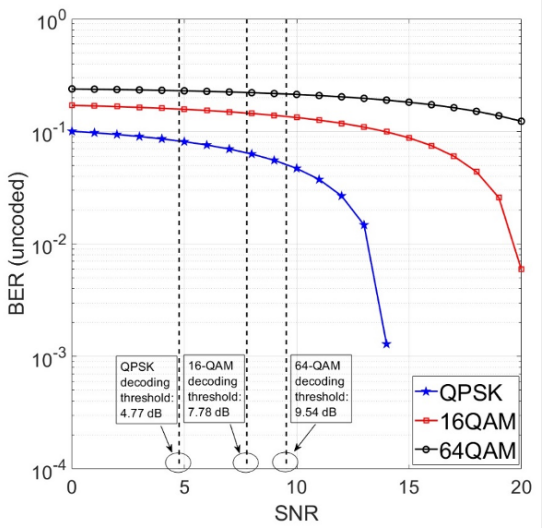
$$G_{M\times M} = \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1M} \\ g_{21} & g_{22} & \cdots & g_{2M} \\ \vdots & \vdots & \vdots & \vdots \\ g_{M1} & g_{M2} & \cdots & g_{MM}\end{bmatrix}. \tag{35}$$

According to the allocation matrix X in equation (7), we define the number of UEs occupying the subchannel c $(1 \leq c \leq R)$ in X as

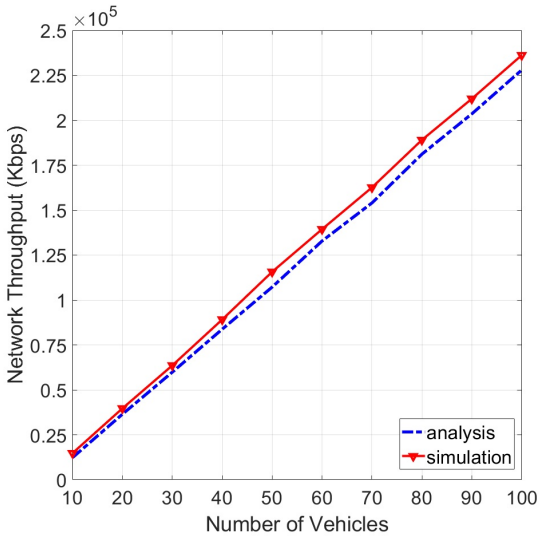
$$sx_c = \sum_{i=1}^{M} x_{ic}, (1 \leq c \leq R). \tag{36}$$

Authorized licensed use limited to: Mississippi State University Libraries. Downloaded on September 11,2023 at 23:37:11 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3294279



(a) Analytical results of BER performance for SC-FDMA over Nakagami fading channels



(b) Network throughput

Fig. 11: Network throughput comparison between analytical and simulation results

In order to obtain interference level at UE i, suppose UE i occupies n consecutive subchannels $CH_r, CH_{r+1}, ..., CH_{r+n-1}$ ($r \geq 1$ and $r+n-1 \leq R$), we define a vector x_{max} from all relevant n columns in matrix X so that it has the largest sum among $sx_r, sx_{r+1}, ..., sx_{r+n-1}$.

$$x_{max} = \underset{r \leq c \leq r+n-1}{\arg\max} \ (sx_c). \tag{37}$$

Therefore, equation (37) is to figure out, among all n subchannels, which one has the largest number of vehicles to use this channel. That is, which subchannel has the most intensive interference. The vector x_{max} shows the situation that which vehicles to use this subchannel. We assume that all VUEs and CUEs have the same transmission power E_s, and sx_t has the largest value among all n vectors (e.g., $sx_r, sx_{r+1}, ..., sx_{r+n-1}$) described in equation (37), so the interference level [3] at node i is

$$\begin{aligned}\Lambda =& E_s G x_{max} \\ =& E_s \begin{bmatrix} g_{11}x_{1t} + g_{12}x_{2t} + ... + g_{1M}x_{Mt} \\ g_{21}x_{1t} + g_{22}x_{2t} + ... + g_{2M}x_{Mt} \\ \vdots \\ g_{M1}x_{1t} + g_{M2}x_{2t} + ... + g_{MM}x_{Mt} \end{bmatrix},\end{aligned} \tag{38}$$

where the vector $x_{max} = (x_{1t}, x_{2t}, ..., x_{Mt})^T$ is the t-th column of the allocation matrix X.

Therefore, considering a sending node v_i, the corresponding $SINR$ at a receiving node v_j that may be exposed to interference caused by multiple other sending nodes is

$$\mathcal{F}_{ij} = \frac{E_s g_{ij}}{\Lambda_j + P_{N0}} (i \neq j), \tag{39}$$

where P_{N0} is the noise power and Λ_j is the power sum of all interference. We assume all VUEs and CUEs have the same transmission power E_s. Since the interference contributions are overlapped, by applying the central limit theorem, we can replace SNR in (26) with $SINR$ $\mathcal{F}_{ij}$ and use (34) to obtain the BER.

To derive the PER, forward error correction (FEC) code needs to be applied. 3GPP has specified Low-density parity-check (LDPC) code for V2X communications with certain typical coding rates, including 1/3, 1/2, and 2/3. Based on the Gaussian approximation (GA), the BER expressions and the thresholds for error-free decoding with LDPC codes over i.i.d. Nakagami-m fading channels are derived [48]. With rate 1/2 (3, 6) regular LDPC code with block size set to 10^4 and fading depth of $m = 3$, we can derive the decoding threshold d_{thld} (i.e., $BER \leq 10^{-7}$) for the case of QPSK, 16-QAM and 64-QAM is 4.77dB, 7.78dB and 9.54dB, respectively, as shown in Fig. 11(a).

Let $\overline{BER}$ denote the BER after applying the above LDPC FEC code, so its approximate expression is

$$\overline{BER} \approx \begin{cases} BER, \text{ if SINR} \leq d_{thld} \\ 10^{-7}. \text{ otherwise} \end{cases} \tag{40}$$

As a result, if the block length is L_b and we assume an independent bit error, the corresponding PER is $PER = 1-(1-\overline{BER})^{L_b}$ for packets no longer than L_b bits. Assume the number of incoming packets at a node v_i is N_i within duration t, the expression $NC_i = N_i(1-PER)$ then denotes the number of successfully received packets, whose size is $PS_i, 0 \leq i \leq NC_i$ in bits. The network throughput becomes:

$$Thpt = \frac{\sum_{i=1}^{M} \sum_{x=1}^{NC_i} PS_x}{t}. \tag{41}$$

To validate the proposed analytical model, both the analytical results and the simulation results are demonstrated in Fig. 11(b). The simulation is performed in NS-3 under the same configuration and parameters. The curve from analysis tightly

[3]Since multiple subchannels may be assigned to a UE, the interference strength for different subchannel may vary. The equation (38) shows the largest interference among all the allocated subchannels of the UE.

Authorized licensed use limited to: Mississippi State University Libraries. Downloaded on September 11,2023 at 23:37:11 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3294279

TABLE IV: Key Parameters in simulation.

Parameter	Value
Frequency/Band width	5.9 GHz/20 MHz
Subcarrier Space/Slot duration	15 kHz/1 ms
Number of UEs/RSUs in one cell	200 to 1200/4
Transmission range	RSU: 200 m CUEs and VUEs: 100 m
Radius of positioning zone Z_r	5 m
Learning rate θ	0.85
Discount rate γ	0.95
Policy adjustment probability ϵ	dynamically from 0.5 to 0.1
Policy adjustment probability η	dynamically from 0.85 to 0.95
LDPC coding rate	0.5
Travel velocity	VUEs: 0 to 50 km/h CUEs: 0 to 10 km/h
Area size/Grid size	1000 m x 1000 m/10x10
Channel fading model	Nakagami fading
Fading depth	0.8 if distance ≤ 80 m 1.5 otherwise
Path loss exponent	3
Defuzzification threshold $defuz$	0.3
Number of RBs per subchannel / Number of subchannels in data channel	5/22
Modulations	16-QAM: safety services 64-QAM: non-safety services

follows the curve from simulation. The simulation setup is elaborated in Section VI. The narrow gap between the two curves mainly stems from the approximate operations, such as BER calculation in (40).

VI. Performance Evaluation

A. Simulation Setup

We simulate a cellular V2X network by building up the aforementioned system model in NS-3 [49]. It comprises one base station, a certain number of UEs, such as CUEs, VUEs and RSUs. In order to reflect how the proposed FAQ deals with the dynamics and high demands of V2X communications in urban regions, in which the centralized resource allocation method support various services, a Manhattan grid is built in SUMO [50], and it is linked with the V2X network in NS-3.

In the simulation, a general road network, the Manhattan Road network with 11 vertical streets and 11 horizontal streets forming a grid with area of 1 km^2 has been developed in SUMO. Two parallel and adjacent streets are placed 100 m apart. To make the mobility of UEs more realistic, we use a dynamic mobility model. In this model, VUEs travel at a speed from 0 to 50 km/h and CUEs have a speed from 0 to 10 km/h. Random trips are set up for each UE. More specifically, UEs start to travel from respective locations initialized randomly, pick up their own routes and go through the Manhattan grid and finally reach their destinations. During the whole process, they experience acceleration, deacceleration and static stages. They may stop at an intersection and wait for the traffic light. The entire procedure is very similar to the real-world situations. When one trip ends, they will start another trip until the end of the simulation. Regarding the multiplexing mode in the simulations, frequency division multiplexing access (FDMA) is leveraged at the physical layer. In this mode, each UE may be assigned different subchannels and subframes to access the data channel.

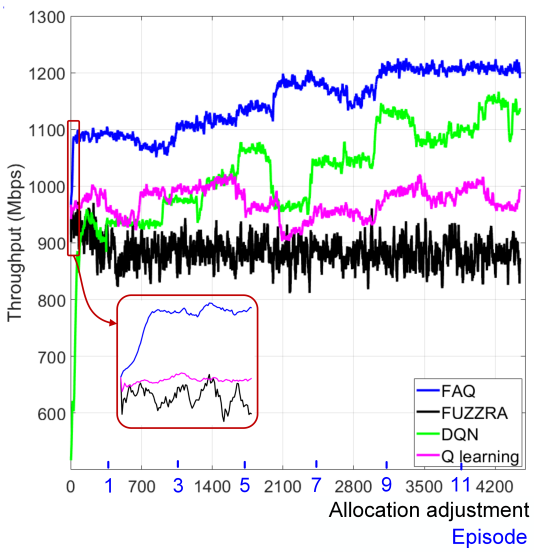


Fig. 12: The network throughput during the process of learning for the case of 200 VUEs in the network.

Table IV captures the simulation parameters. In the resource pool, there are 20 subframes and 22 subchannels. To avoid large delays, a transmission request should be approved in the next T_{sf} subframes if suitable resources can be located. The value of T_{sf} depends on the requirement of data rate for a particular V2X service. For instance, regarding cooperative manoeuvre (SER2), we have $T_{sf} = 20$, which means a message from *SER2* should be sent no later than 20 ms since the corresponding request has been received. Otherwise, it would be declined for transmission.

In the simulation, the proposed learning is carried out according to Algorithm 1, and the allocated resources will be re-adjusted on a one-by-one basis during the learning. Since there are a large number of allocations, each iteration consists of all adjustments for each allocation. The number of allocations is contingent of UEs' volume in the network and different communication situations. For instance, all VUEs periodically broadcast *SER1* messages, while the communications of two non-safety services and *SER3* only take place when necessary, according to their respective definitions. The learning process continues until a convergence of network throughput is observed.

B. Simulation Results and Discussions

The performance of the proposed FAQ algorithm has been assessed by extensive simulations with the configurations and parameters elaborated above. With the same configurations, other three advanced resource allocation algorithms have been simulated as benchmarks. One is the conventional Q-learning. It explores all possible allocations randomly to figure out the best one. The other one is fuzzy logic based resource allocation FUZZRA [25], which merely relies on a fuzzy-logic based model for the resource allocation. The third one is a deep Q learning model (DQN). It consists of a V2X network in NS-3, an agent in OpenAI Gym [51] realized by a deep neural network that mapping states to Q values for different actions, and an interface (e.g., NS3-gym [52]) connecting the

Authorized licensed use limited to: Mississippi State University Libraries. Downloaded on September 11,2023 at 23:37:11 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3294279

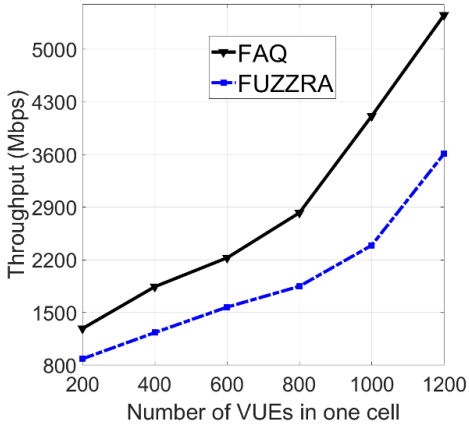


Fig. 13: Network throughput comparison between two resource allocation schemes with different number of VUEs in the network.

V2X network to the agent. The deep neural network has a construction of one input layer, 8 hidden layers and one output layer.

Fig. 12 compares the network throughput of the four allocation methods. The proposed FAQ outperforms the other three peers in terms of network throughput. Moreover, the corresponding curve quickly converges after 9 episodes of learning and maintains a higher level than the other three through the whole process. At episode 9, its stable throughput is 10%, 20% and 33% higher than the DQN, the Q learning and the FUZZRA, respectively. It is worth noting that each episode means accomplishing adjustments for all requests. In this simulation, one episode has 350 allocation adjustments as there are 350 requests from all UEs. By contrast, although the DQN performs better than the Q-learning after episode 5, both of them fail to reach a stable throughput within 13 episodes, and have a lower throughput than the FAQ. The throughput standing for the FUZZRA always keeps steady but with fluctuations during the process, because the FUZZRA reconsiders the allocations each time using the same logic without any progress resulted from learning like the other three.

To better understand the the scalability of the FAQ, various densities of VUEs in an urban area are studied. The number of vehicles ranges from 200 to 1200, which covers the situations from the sparse to the dense. Fig. 13 shows how the network throughput changes for both FAQ and FUZZRA. As the density of UEs increases, more transmission requests have been sent to the base station. The network throughput is much higher if the FAQ is applied to the V2X communications. The gap of throughput between two algorithms is widened as the number of UEs increases in the network. This implies that the proposed FAQ can reuse resources more effectively yet without affecting communications. The results representing the DQN and the Q-learning are not included as they cannot reach a stable stage within the same period of learning.

The average packet delivery ratio is also optimized by the FAQ as the interference is substantially suppressed, as shown in Fig. 14. The PDR curves for the FAQ, the DQN and the Q learning rise during the learning process. However, only

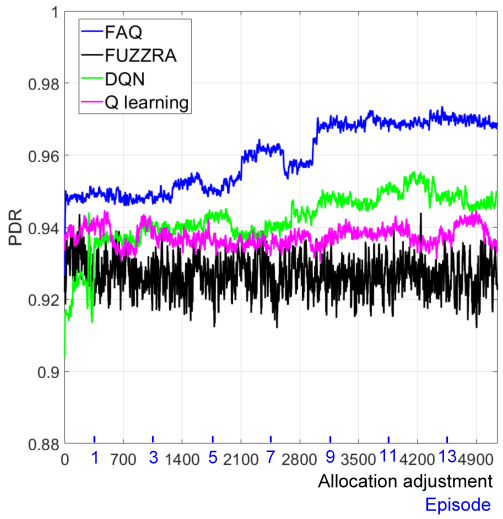


Fig. 14: Packet delivery ratio (PDR) comparison during the learning process.

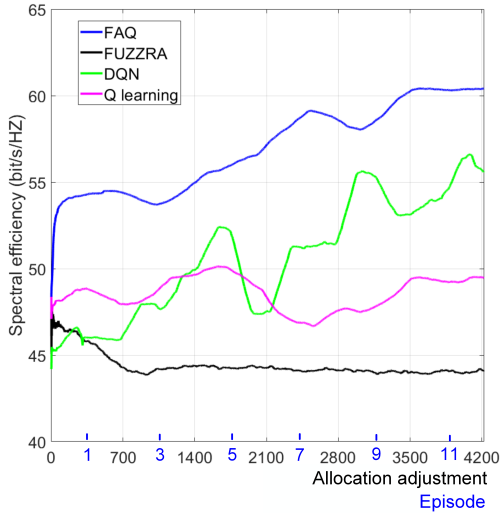


Fig. 15: Spectral efficiency of the network during the learning process.

the FAQ curve converges and reaches a stable values of 97% after episode 9. The other two curves do not converge as fast as the FAQ one, which has a similar situation in Fig. 12. The fuzzy-logic assisted action selector in the FAQ takes into account the interference factor to mitigate the interference while increasing the spectrum reuse among UEs in non-conflicting zones. Additionally, the FUZZRA curve maintain almost the same level during the observation period.

To improve the frequency utilization and satisfy as many as requests from vehicles, channels are reused to accommodate as many concurrent transmissions as possible. We use network spectral efficiency to evaluate the overall information bits being transferred per second per Hertz in the network. The total bandwidth of the data channel is 20 MHz. Although all the four allocation schemes start from an approximate value, Fig. 15 shows that the FAQ has the highest spectral efficiency and its curve converges rapidly at episode 9. Again, the DQN and the Q learning do not converge at that time, and they have lower spectral efficiency. Additionally, the FUZZRA

Authorized licensed use limited to: Mississippi State University Libraries. Downloaded on September 11,2023 at 23:37:11 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3294279

curve descends slightly over the observation period, which shares the same trend in the network throughput and the PDR simulations.

C. Computational Complexity Analysis

The computational complexity of the proposed FAQ is $O(N^3)$, and the proof is elaborated in Appendix C.

VII. Conclusion

In this study, we propose a novel reinforcement learning model to tackle how to allocate the limited resources in V2X communications within the context of more demanding V2X services being deployed. These services, on the one hand, consume considerable resources, and on the other hand, have higher requirements. The high density of vehicles in urban areas even makes the problem more challenging. Based on such observations, a fuzzy-logic assisted Q learning model is developed to dynamically allocate resources. Since there are infinite possibilities in choosing resources for requests, the proposed FAQ strategy aims to maximize the network throughput while considering inequalities among different V2X services. The integration of fuzzy logic into Q learning not only can significantly accelerate the learning process, but also improve the learning outcome, because the fuzzy-logic component can predict instant rewards for choosing an action, and reinforcement learning considers long-term rewards of an allocation. Through extensive simulations in NS-3, the results demonstrate the advantage of the proposed algorithm over other alternatives. Additionally, a mathematical model is built to analyze the network throughput based on the allocation matrix and derivation of packet error rate. In the 6G era, UAVs may play a key role and extend the existing ground-based V2X communications to more comprehensive three-dimensional (3D) situations, which may involve the coexistence of vehicle-to-UAV, UAV-to-UAV and V2V communications. The authors will study how to model the channels, allocate resources and improve the spectral efficiency in such scenarios, in which UAVs flying within air corridors and vehicles moving on roads interact with each other.

Acknowledgment

This work is supported in part by National Science Foundation of US (Grant No. 2120442), and it is also supported in part by the Natural Science Research Start-up Foundation of Recruiting Talents of Nanjing University of Posts and Telecommunications (Grant No. NY222102), the National Natural Science Foundation of P.R. China (Grant No. 62272244).

Appendix

A. Action Space

The base station responds to the requests from UEs. Each type of service request needs a different number of successive subchannels and different latency. BW determines the required number of successive subchannels. different types of requests have different BWs, as illustrated in Table II. Although a request may be assigned a number of consecutive subchannels starting from any subchannel, the actual starting subchannel of a request is designed to be aligned with the one dedicated types of request (i.e., aligned with requests of SER1 that need the minimum number of subchannels for each allocation). Such an observation accords with the interference caused by the inevitable channel reuse due to the limited resource, compared to the numerous requests from UEs. For instance, a request for SER3 needs 18 subchannels. There are only two possible allocations at one subframe: subchannels ranging from 1 to 18, and subchannels ranging from 4 to 21. That is, we do not consider allocating subchannels from 2 to 19 or from 3 to 20, because subchannels 1 to 3 have or will be likely assigned to a SER1 request, which will result in severe interference. Thus, the starting subchannel for a request could be only chosen from subchannel 1, 4, 7, 10 or 13. Fig. 16 lists all possible allocations for the six types of services in a subframe.

- SER1 type of requests requiring BW=3 can be assigned 7 possible subchannel ranges (i.e., [1 $\sim$ 3], [4 $\sim$ 6], [7 $\sim$ 9], [10 $\sim$ 12], [13 $\sim$ 15], [16 $\sim$ 18] and [19 $\sim$ 21]).
- SER2 type of requests requiring BW=10 can be assigned 5 possible subchannel ranges (i.e., [1 $\sim$ 10], [4 $\sim$ 13], [7 $\sim$ 16], [10 $\sim$ 19] and [13 $\sim$ 22]).
- SER3 type of requests requiring BW=18 can be assigned 2 possible subchannel ranges (i.e., [1 $\sim$ 18], [4 $\sim$ 21]).
- SER4 type of requests requiring BW=8 can be assigned 5 possible subchannel ranges (i.e., [1 $\sim$ 8], [4 $\sim$ 11], [7 $\sim$ 14], [10 $\sim$ 17], and [13 $\sim$ 20]).
- SER5 and SER6 type of requests requiring BW=12 can be assigned 4 possible subchannel ranges (e.g., [1 $\sim$ 12], [4 $\sim$ 15], [7 $\sim$ 18], and [10 $\sim$ 21]).

In this study, we consider the allocation time slot as 20 milliseconds (20 subframes). For a request with a maximum delivery latency of 10 ms, the base station will assign it with a number of subchannels within the 10 subframes. That is, the base station will choose resources from 7*20=140 combinations of subchannels and subframes.

B. Details of the deployed V2X services

We deploy six different V2X services in the network. They include four safety-related applications and two infotainment applications. The corresponding message types, transmission frequencies, and data rates are shown in Table V.

C. Complexity of FAQ

The computational complexity of the FAQ is $O(N^3)$, as shown below.

Proof: We first define that $\mathbb{N}_{req} = \{rq_1, rq_2, ...\}$ is the set of transmission requests from all M UEs, and $N_q = |\mathbb{N}_{req}|$ is the total number of requests. The FAQ algorithm will try to assign resources for each request. There are three options to choose an action for each request in the FAQ: using the fuzzy logic, the maximum Q-value or the random selection. The corresponding probability of utilizing the three methods is ρ_1, ρ_2 and ρ_3, respectively, and $\rho_1 + \rho_2 + \rho_3 = 1$. For $\forall rq_i \in \mathbb{N}_{req}$, if the fuzzy-logic method is used, according to [25], the computing time is $K_c N_q$ for choosing an action, where

Authorized licensed use limited to: Mississippi State University Libraries. Downloaded on September 11,2023 at 23:37:11 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3294279

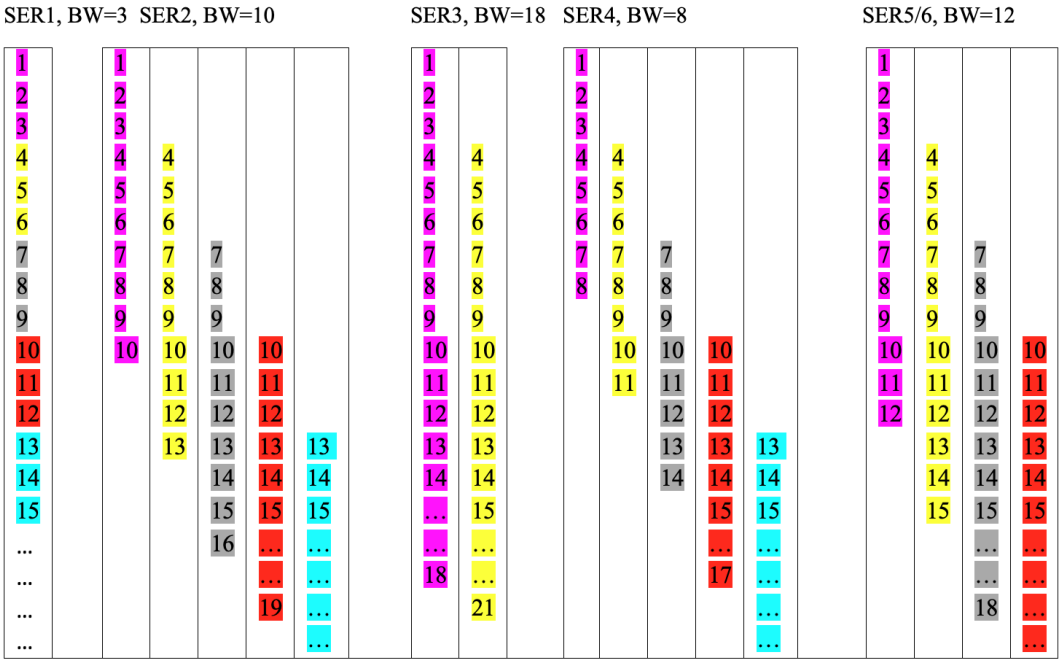


Fig. 16: All possible subchannel assignments for different types of requests

TABLE V: Type of services implemented in the V2X communications.

Deployed Services	Message Type	Beacon Frequency	End-to-End latency	Typical Data Rate	Packet size (Bytes)
Cooperative awareness (SER1)	Periodic broadcast	100 ms	100 ms	5-100 Kbps	300
Cooperative manoeuvre (SER2)	Periodic broadcast	10 ms	20 ms	2-5 Mbps	1500
Enhanced sensing (SER3)	Event-driven Broadcast	N/A	20 ms	10-25 Mbps	3000
Dynamic traffic control and warning (SER4)	Periodic broadcast	500 ms	500 ms	0.5-2 Mbps	1200
Non-safety non-real-time content (SER5)	Non-periodic unicast	N/A	300-65535 ms	1-5 Mbps	1500
Non-safety real-time content (SER6)	Non-periodic unicast	N/A	20 ms	5-10 Mbps	1500

K_c is a constant. If the maximum Q-value method is utilized, the execution times is $|A|$, where $|A|$ is the total number of actions in the action space A. Likewise, the computation is a constant T_{rdm} for the random selection method. Therefore, the execution time of an action selection for an arbitrary request rq_i is:

$$T_{rq} = \rho_1 K_c N_q + \rho_2 |A| + \rho_3 T_{rdm}. \tag{42}$$

The steps from line 6 to line 17 in Algorithm1 is repeatedly executed $N_q T_{rq}$ times. Furthermore, according to study [53], the proposed FAQ learning reaches a goal state at most $O(|A|N_q)$ steps. Finally, the total number of executions is acquired:

$$\begin{aligned} T_{FAQ} &= |A| N_q (N_q T_{rq}) \\ &= \rho_1 |A| K_c N_q^3 + (\rho_2 |A|^2 + \rho_3 |A| T_{rdm}) N_q^2. \end{aligned} \tag{43}$$

From the above, the computational complexity of the FAQ is $O(N^3 + N^2)$, since $|A|, K_c, T_{rdm}, \rho_1, \rho_2$ and ρ_3 are all constants. The expression $O(N^3 + N^2)$ can further be deemed as $O(N^3)$.

REFERENCES

[1] 3GPP, "TS 22.185: Service requirements for V2X services; Stage 1 (v15.0.0, Release 15)," 3GPP, Tech. Rep. 36.213, June 2017.

[2] 3GPP, "TR 22.886: Study on enhancement of 3GPP Support for 5G V2X Services (v16.2.0, Release 16)," 3GPP Tech. Rep., Dec. 2018.

[3] 3GPP, "TR 36.885 Study on LTE-based V2X services (v14.0.0, Release 14)," 3GPP, Tech. Rep. 36.885, July 2016.

[4] W. Saad, M. Bennis, and M. Chen, "A vision of 6G wireless systems: Applications, trends, technologies, and open research problems," *IEEE Network*, vol. 34, no. 3, pp. 134–142, 2020.

[5] Y. Zhu, B. Mao, and N. Kato, "Intelligent reflecting surface in 6G vehicular communications: A survey," *IEEE Open Journal of Vehicular Technology*, vol. 3, pp. 266–277, 2022.

[6] Y. Zhu, B. Mao, Y. Kawamoto, and N. Kato, "Intelligent reflecting surface-aided vehicular networks toward 6G: Vision, proposal, and future directions," *IEEE Veh. Technol. Mag.*, vol. 16, no. 4, pp. 48–56, 2021.

[7] M. Noor-A-Rahim, Z. Liu, H. Lee, M. O. Khyam, J. He, D. Pesch, K. Moessner, W. Saad, and H. V. Poor, "6G for vehicle-to-everything (V2X) communications: Enabling technologies, challenges, and opportunities," *Proceedings of the IEEE*, vol. 110, no. 6, pp. 712–734, 2022.

[8] A. Mekrache, A. Bradai, E. Moulay, and S. Dawaliby, "Deep reinforcement learning techniques for vehicular networks: Recent advances and future trends towards 6G," *Veh. Commun.*, vol. 33, p. 100398, 2022.

[9] F. Tariq, M. R. A. Khandaker, K. Wong, M. A. Imran, M. Bennis, and M. Debbah, "A speculative study on 6G," *IEEE Wirel. Commun.*, vol. 27, no. 4, pp. 118–125, 2020.

[10] H. Ouamna, Z. Madini, and Y. Zouine, "6G and V2X communications: Applications, features, and challenges," in *2022 8th International Conference on Optimization and Applications (ICOA)*, 2022, pp. 1–6.

[11] R. Molina-Masegosa and J. Gozalvez, "LTE-V for sidelink 5G V2X vehicular communications: A new 5G technology for short-range vehicle-to-everything communications," *IEEE Vehicular Technology Magazine*, vol. 12, no. 4, pp. 30–39, 2017.

[12] A. Dayal, V. K. Shah, B. Choudhury, V. Marojevic, C. Dietrich, and J. H. Reed, "Adaptive semi-persistent scheduling for enhanced on-road safety in decentralized V2X networks," in *2021 IFIP Networking Conference (IFIP Networking)*, 2021, pp. 1–9.

[13] B. Toghi, M. Saifuddin, H. N. Mahjoub, M. O. Mughal, Y. P. Fallah, J. Rao, and S. Das, "Multiple access in cellular V2X: performance

Authorized licensed use limited to: Mississippi State University Libraries. Downloaded on September 11,2023 at 23:37:11 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3294279

analysis in highly congested vehicular networks," in *2018 IEEE Vehicular Networking Conference, VNC 2018, Taipei, Taiwan, December 5-7, 2018*. IEEE, 2018, pp. 1–8.

[14] Y. Liang, X. Chen, S. Chen, and Y. Chen, "Cooperative resource sharing strategy with eMBB cellular and C-V2X slices," in *26th IEEE International Conference on Parallel and Distributed Systems (ICPADS), Hong Kong, 2020.* IEEE, 2020.

[15] S. Yi, G. Sun, and X. Wang, "Enhanced resource allocation for 5G V2X in congested smart intersection," in *92nd IEEE Vehicular Technology Conference, VTC Fall 2020, Victoria, BC, Canada, November 18 - December 16, 2020*. IEEE, 2020, pp. 1–5.

[16] L. Liang, H. Ye, and G. Y. Li, "Spectrum sharing in vehicular networks based on multi-agent reinforcement learning," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 10, pp. 2282–2292, 2019.

[17] 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, 2021.

[18] S. Kim, B.-J. Kim, and B. B. Park, "Environment-adaptive multiple access for distributed V2X network: A reinforcement learning framework," in *2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring)*. IEEE, 2021, pp. 1–7.

[19] N. Bonjorn, F. Foukalas, F. Canellas, and P. Pop, "Cooperative resource allocation and scheduling for 5G eV2X services," *IEEE Access*, vol. 7, pp. 58 212–58 220, 2019.

[20] X. Zhang, M. Peng, S. Yan, and Y. Sun, "Deep-reinforcement-learning-based mode selection and resource allocation for cellular V2X communications," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 6380–6391, 2020.

[21] F. Tang, Y. Zhou, and N. Kato, "Deep reinforcement learning for dynamic uplink/downlink resource allocation in high mobility 5G hetnet," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 12, pp. 2773–2782, 2020.

[22] M. Zhang, A. Kumar, P. H. J. Chong, H. C. B. Chan, and B. Seet, "Resource allocation based performance analysis for 5G vehicular networks in urban areas," in *2020 IEEE Wireless Communications and Networking Conference Workshops, WCNC Workshops 2020, Seoul, Korea (South), April 6-9, 2020*. IEEE, 2020, pp. 1–6.

[23] D. Sempere-García, M. Sepulcre, and J. Gozalvez, "LTE-V2X mode 3 scheduling based on adaptive spatial reuse of radio resources," *Ad Hoc Networks*, vol. 113, p. 102351, 2021.

[24] G. Nardini, A. Virdis, C. Campolo, A. Molinaro, and G. Stea, "Cellular-V2X communications for platooning: Design and evaluation," *Sensors*, vol. 18, no. 5, p. 1527, 2018.

[25] M. Zhang, Y. Dou, P. H. J. Chong, H. C. B. Chan, and B. Seet, "Fuzzy logic-based resource allocation algorithm for V2X communications in 5g cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 8, pp. 2501–2513, 2021.

[26] 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 Trans. Veh. Technol.*, vol. 71, no. 11, pp. 12 044–12 056, 2022.

[27] M. Yuan, Q. Cao, M.-O. Pun, Y. Chen *et al.*, "Fairness-oriented user scheduling for bursty downlink transmission using multi-agent reinforcement learning," *APSIPA Transactions on Signal and Information Processing*, vol. 11, no. 1, 2022.

[28] D. Zavyalova and V. Drozdova, "5G scheduling using reinforcement learning," in *2020 International Multi-Conference on Industrial Engineering and Modern Technologies (FarEastCon)*, 2020, pp. 1–5.

[29] T. Kim and M. Min, "A low-complexity algorithm for a reinforcement learning-based channel estimator for MIMO systems," *Sensors*, vol. 22, no. 12, p. 4379, 2022.

[30] M. S. Oh, S. Hosseinalipour, T. Kim, C. G. Brinton, and D. J. Love, "Channel estimation via successive denoising in MIMO OFDM systems: A reinforcement learning approach," in *ICC 2021 - IEEE International Conference on Communications, Montreal, QC, Canada, June 14-23, 2021*. IEEE, 2021, pp. 1–6.

[31] L. F. Abanto-Leon, A. Koppelaar, and S. M. H. de Groot, "Subchannel allocation for vehicle-to-vehicle broadcast communications in mode-3," in *2018 IEEE Wireless Communications and Networking Conference, WCNC 2018, Barcelona, Spain, April 15-18, 2018*. IEEE, 2018, pp. 1–6.

[32] L. F. Abanto-Leon, A. Koppelaar, and S. Heemstra de Groot, "Network-assisted resource allocation with quality and conflict constraints for V2V communications," in *2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*, 2018, pp. 1–5.

[33] K. Sehla, T. M. T. Nguyen, G. Pujolle, and P. B. Velloso, "A new clustering-based radio resource allocation scheme for C-V2X," in *2021 Wireless Days, WD 2021, Paris, France, June 30 - July 2, 2021*. IEEE, 2021, pp. 1–8.

[34] G. Cecchini, A. Bazzi, B. M. Masini, and A. Zanella, "Localization-based resource selection schemes for network-controlled LTE-V2V," in *2017 International Symposium on Wireless Communication Systems (ISWCS)*, 2017, pp. 396–401.

[35] R. Fritzsche and A. Festag, "Location-based scheduling for cellular V2V systems in highway scenarios," in *2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*. IEEE, 2018, pp. 1–5.

[36] M. Gonzalez-Martin, M. Sepulcre, R. Molina-Masegosa, and J. Gozálvez, "Analytical models of the performance of C-V2X mode 4 vehicular communications," *IEEE Trans. Veh. Technol.*, vol. 68, no. 2, pp. 1155–1166, 2019.

[37] C. Wu, T. Yoshinaga, Y. Ji, T. Murase, and Y. Zhang, "A reinforcement learning-based data storage scheme for vehicular ad hoc networks," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 7, pp. 6336–6348, 2016.

[38] S. Vemireddy and R. R. Rout, "Fuzzy reinforcement learning for energy efficient task offloading in vehicular fog computing," *Computer Networks*, vol. 199, p. 108463, 2021.

[39] I. Kilanioti, G. Rizzo, B. M. Masini, A. Bazzi, D. P. M. Osorio, F. Linsalata, M. Magarini, D. Löschenbrand, T. Zemen, and A. Kliks, "Intelligent transportation systems in the context of 5G-Beyond and 6G Networks," in *2022 IEEE Conference on Standards for Communications and Networking (CSCN)*, 2022, pp. 82–88.

[40] Z. Zhang, Y. Xiao, Z. Ma, M. Xiao, Z. Ding, X. Lei, G. K. Karagiannidis, and P. Fan, "6G wireless networks: Vision, requirements, architecture, and key technologies," *IEEE Vehicular Technology Magazine*, vol. 14, no. 3, pp. 28–41, 2019.

[41] A. Bazzi, C. Campolo, V. Todisco, S. Bartoletti, N. Decarli, A. Molinaro, A. O. Berthet, and R. A. Stirling-Gallacher, "Toward 6G vehicle-to-everything sidelink: Nonorthogonal multiple access in the autonomous mode," *IEEE Vehicular Technology Magazine*, vol. 18, no. 2, pp. 50–59, 2023.

[42] Z. Lu, T. Zhang, X. Ji, B. Qian, L. Jiao, and H. Zhou, "Personalized wireless resource allocation in multi-connectivity B5G C-V2X networks," in *2022 14th International Conference on Wireless Communications and Signal Processing (WCSP)*, 2022, pp. 1–6.

[43] S. B. Prathiba, G. Raja, S. Anbalagan, K. Dev, S. Gurumoorthy, and A. P. Sankaran, "Federated learning empowered computation offloading and resource management in 6G-V2X," *IEEE Transactions on Network Science and Engineering*, vol. 9, no. 5, pp. 3234–3243, 2022.

[44] M. Döttling, W. Mohr, and A. Osseiran, *Radio technologies and concepts for IMT-advanced*. John Wiley & Sons, 2009.

[45] 3GPP, "Study on evaluation methodology of new vehicle-to- everything (V2X) use cases for LTE and NR," 3GPP, Tech. Rep. TR37.885, June 2019.

[46] J. J. Sanchez-Sanchez, M. C. Aguayo-Torres, and U. Fernandez-Plazaola, "BER analysis for zero-forcing SC-FDMA over nakagami-m fading channels," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 8, pp. 4077–4081, 2011.

[47] F. J. López-Martínez, E. Martos-Naya, J. F. Paris, and U. Fernández-Plazaola, "Generalized BER analysis of qam and its application to mrc under imperfect CSI and interference in ricean fading channels," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 5, pp. 2598–2604, 2010.

[48] B. S. Tan, K. H. Li, and K. C. Teh, "Performance analysis of LDPC codes with maximum-ratio combining cascaded with selection combining over nakagami-m fading," *IEEE Transactions on Wireless Communications*, vol. 10, no. 6, pp. 1886–1894, 2011.

[49] G. F. Riley and T. R. Henderson, "The NS-3 network simulator," in *Modeling and Tools for Network Simulation*. Springer, 2010, pp. 15–34.

[50] P. A. Lopez, M. Behrisch, L. Bieker-Walz, J. Erdmann, Y.-P. Flötteröd, R. Hilbrich, L. Lücken, J. Rummel, P. Wagner, and E. WieBner, "Microscopic traffic simulation using SUMO," in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2018, pp. 2575–2582.

[51] G. Brockman, V. Cheung, L. Pettersson, J. Schneider, J. Schulman, J. Tang, and W. Zaremba, "OpenAI Gym," 2016, arXiv:1606.01540.

[52] P. Gawłowicz and A. Zubow, "NS-3 meets OpenAI Gym: The playground for machine learning in networking research," ser. MSWIM '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 113–120. [Online]. Available: https://doi.org/10.1145/3345768.3355908

Authorized licensed use limited to: Mississippi State University Libraries. Downloaded on September 11,2023 at 23:37:11 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3294279

[53] S. Koenig and R. G. Simmons, "Complexity analysis of real-time reinforcement learning," in *Proceedings of the 11th National Conference on Artificial Intelligence. Washington, DC, USA, July 11-15, 1993*, R. Fikes and W. G. Lehnert, Eds. AAAI Press / The MIT Press, 1993, pp. 99–107.

Authorized licensed use limited to: Mississippi State University Libraries. Downloaded on September 11,2023 at 23:37:11 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.