UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

DOCTORAL THESIS

# Radio Resource Management for V2X in Cellular systems

*Author*: Haider Daami R. Albonda

*Supervisor:* Dr. Jordi Pérez-Romero

*A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy in the Mobile Communications Research Group Department of Signal Theory and Communications*

January 2020

# *Abstract*

Emerging 5G wireless technology will support various services and use cases with vastly heterogeneous requirements, such as enhanced mobile broadband (eMBB), massive machine-type communications (mMTC) and ultra-reliable and low-latency communications (uRLLC), which include e.g., industry automation and some vehicle-to-everything (V2X) services. In order to support these heterogeneous requirements, the 5G system will rely on network slicing as one of the cornerstone features. Network slicing allows composing multiple dedicated logical networks with specific functionality running on top of a common infrastructure. At the radio access network (RAN), the use of network slicing involves the assignment of radio resources to each slice in accordance with its expected requirements and functionalities.

In this context, this thesis has focused on the provision of cellular V2X communications, which have attracted great interest for 5G due to the potential of improving traffic safety and enabling new services related to intelligent transportation systems. These types of services have strict requirements on reliability, access availability, and end-to-end (E2E) latency. V2X requires advanced network management techniques that must be developed based on the characteristics of the networks and traffic requirements. One of the new features in the cellular networks for supporting V2X is the Sidelink (SL), which enables the direct communication between vehicles without passing through the base station. As a result, vehicle-to-vehicle (V2V) communications can be performed using two communication modes, namely cellular communication based on uplink/downlink and direct V2V communications using sidelink. By properly selecting the operating mode according to the requirements of services, significant benefits such as efficient utilization of the spectrum, low latency, high reliability, efficiency and scalability can be achieved. However, selecting the appropriate operating mode according to the requirements of V2V services and allocating the radio resources to V2V communications becomes a challenging issue. Hence in

this thesis, we address some of the challenges brought by the integration of V2V communication in cellular systems, and validate the potential of this technology by means of proper resource management solutions. Our main contributions lie in the context of RAN slicing, mode selection and radio resource management mechanisms.

First, we propose an efficient RAN slicing scheme for supporting the provision of V2X and eMBB communications. It is based on an off-line reinforcement learning followed by a low-complexity heuristic algorithm, which allocates radio resources to different slices with the target of maximizing the resource utilization while ensuring the availability of resources to fulfill the requirements of the traffic of each RAN slice. A simulation-based analysis is presented to assess the performance of the proposed solution. The simulation results have shown that the proposed algorithm improves the network performance in terms of resource utilization, latency, achievable data rate, and outage probability.

Second, we investigate the mode selection for V2Vcommunication in a cellular system, in order to decide whether to use the sidelink or the cellular mode. Specifically, we propose two different mode selection strategies that take into account the quality of the links between V2V users in sidelink mode and between the base station and vehicles in cellular mode, the available resources, and the network traffic load situation. Our proposed mode selection strategies can bring significant gains to V2V-enabled networks in terms of different performance metrics such as resource utilization, latency, achievable data rate, and outage probability.

Third, we propose a novel joint mode selection and radio resource allocation strategy for V2V communications with the target of minimizing the probability of exceeding the maximum delay under the constraint of satisfying the reliability requirement. It includes a centralized and a distributed scheduling strategy for cellular and sidelink modes, respectively. The efficiency of the proposed schemes has been demonstrated through simulations and comparisons with benchmark schemes. The proposed mode selection and radio resource allocation strategies achieved significant improvements in the network performance in terms of latency of V2V services, packet success rate and resource utilization.

# *Acknowledgments*

The underlying thesis would not have been possible to complete without receiving the support and help of many people. First and foremost, I would like to sincerely thank my advisor Dr. Jordi Pérez-Romero for the continuous support of my Ph.D study, for his patience, motivation, and immense knowledge. It is an honor for me to be his Ph.D. student. I appreciate all his contributions of time, ideas, and efforts to make my Ph.D. experience productive and stimulating. Furthermore I would like to thanks the members of Mobile Communications Research Group for the exceptional experience among them as well as good advice and collaboration.

I would also like to thank the Iraqi ministry of higher education and scientific research, and the Spanish Research Council for the funding sources that made my Ph.D. work possible. I would certainly like to thank all my friends for giving me the courage throughout these three years of PhD. You have greatly simplified the accomplishment of this challenging mission.

Finally, my deepest gratitude goes to my parents, my brothers and my sisters for their prayers and infinite support in every single aspect of my life. Words are unable to express how I am indebted to them in my life.

# Contents

# List of Figures

# List of Tables

# Abbreviations

| | |
|---|---|
| 3GPP | Third Generation Partnership Project |
| 3G | Third-Generation |
| 4G | Fourth-Generation |
| 5G | Fifth Generation |
| AF | Application Function |
| AMF | Access and Mobility Function |
| AN | Access network |
| AUSF | Authentication Server Function |
| BLMAB | Budgeted Lock-up Multi Armed Problem |
| CAM | Cooperative Awareness Message |
| CN | Core network |
| C-V2X | Cellular V2X |
| CP | Control plane |
| CSI | Channel State Information |
| CUs | Cellular users |
| D2D | Device-to-Device |
| DL | Downlink |
| DENM | Decentralized Environmental Notification Message |
| DSRC | Dedicated Short-Range Communications |
| eNB | E-UTRAN Node B |
| eMBB | Enhanced mobile broadband |
| eMBMS | Evolved multimedia broadcast multicast services |
| EPC | Evolved Packet Core |

| | |
|---|---|
| ETSI | European Telecommunications Standards Institute |
| eV2X | Enhanced V2X |
| E-UTRAN | Evolved Universal Terrestrial Radio Access Network |
| GBR | Guaranteed Bit Rate |
| ITS | Intelligent Transport Systems |
| LTE | Long Term Evolution |
| MBB | Mobile broadband |
| MBMS | Multimedia Broadcast Multicast Services |
| MDP | Markov Decision Process |
| mMTC | Massive machine-type communications |
| NOS | Non-Orthogonal Sharing approach |
| OFDMA | Orthogonal frequency-division multiple access |
| OS | Orthogonal Sharing approach |
| PCF | Policy Control Function |
| PLMN | Public land mobile network |
| ProSe | Proximity Services |
| QoS | Quality of Service |
| RAN | Radio Access Network |
| RB | Resource Block |
| RL | Reinforcement learning |
| RRM | Radio Resource Management |
| RSU | Roadside Unit |
| SCI | Sidelink Control Information |
| SINRs | Signal-to-interference plus-noise ratios |
| SDN | Software Defined Networking |
| SMF | Session Management Function |
| SL | Sidelink |
| SPS | Semi-Persistent Scheduling |

| | |
|---|---|
| TTI | Transmission Time Interval |
| TN | Transport network |
| UE | User Equipment |
| UL | Uplink |
| UP | User plane |
| UDM | Unified Data Management |
| UDR | Unified Data Repository |
| UPF | User Plane Function |
| uRLLC | Ultra-reliable and low-latency communication |
| USDs | User service descriptions |
| V2I | Vehicle-to-Infrastructure |
| V2N | Vehicle-to-Network |
| V2P | Vehicle-to-Pedestrian |
| V2V | Vehicle-to-Vehicle |
| V2X | Vehicle-to-Everything |
| NFV | Network functions virtualization |
| VRU | Vulnerable Road User |

# Chapter 1

# Introduction

## 1.1 Evolution towards 5G networks

The growth of mobile broadband traffic, the number of connected devices, the demand for high data rates, as well as the emergence of various applications are driving the search for exploring new architectures and paradigms to meet the requirements of future cellular networks. Current mobile networks (e.g., Third-Generation (3G), Fourth-Generation (4G), Wi-Fi…etc.) support high spectral efficiency, high data rates, efficient and high speed mobility support, and security mechanisms for mobile broadband (MBB) services. Despite its potential applicability, current cellular systems were originally designed for mobile broadband traffic which has quite different properties and requirements than future applications and are not suitable to satisfy the diverse and unprecedented future service demands such as enhanced mobile broadband (eMBB), ultra-reliable and low-latency communication (uRLLC), and massive machine-type communications (mMTC) service [1-3]. In order to achieve higher performance and different vertical industrial services and new capabilities, Fifth-Generation (5G) technology of cellular networks is expected to satisfy end to end (E2E) quality of service (QoS) in a wider range of use cases such as those dealing with e-health, industry 4.0, vehicular communications, etc. Figure 1.1 provides some of the network requirements to be addressed, alongside some of the developments that are likely to feature as 4G networks evolve into 5G.

The main expectations for the 5G networks are higher performance by achieving new capabilities with high peak data rates (10-20 Gbps), peak spectral efficiency (15-30 bit/s/Hz), high reliability (up to $10^{-6}$), and Low latency (1 to 10 ms) [4]. According to the International Telecommunication

Union (ITU) [4-6], 5G wireless systems will support three generic services, which are classified as eMBB service, uRLLC service, and mMTC service. A succinct characterization of these services can be put forward as follows: (a) eMBB service aims to meet the user's demand for increasingly digital lifestyle stable connections with very high peak data rates and focuses on services that have high requirements for bandwidth such as high definition (HD) videos; virtual reality (VR) ; and augmented reality (AR) ; (b) mMTC supports a massive number of Internet of Things (IoT) devices, which are only sporadically active and send small data payloads; (c) the uRLLC service that are extremely sensitive to the Reliability and latency, such as industry automation and some vehicle-to-everything (V2X) communication use cases ( i.e., lane merging, and platooning uses cases). Various 5G use case scenarios summarized in figure 1.2. 5G intends to guarantee Ultra-low latency (few milliseconds), high reliability (near 100%), and high data rate required in most of the V2X applications which are the focus of this thesis. In response to the requirements of different services and applications, the 5G system aims to provide a flexible platform by deploying advanced heterogeneous network and slicing technologies.

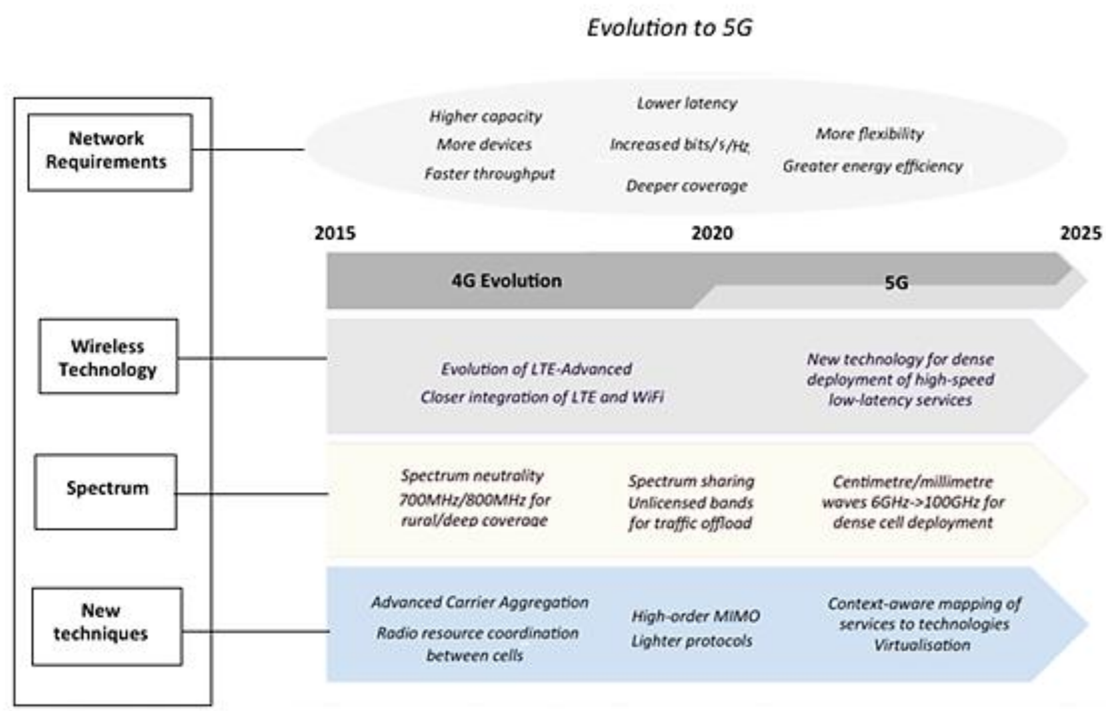


**Figure 1.1**: Evolution towards 5G systems.

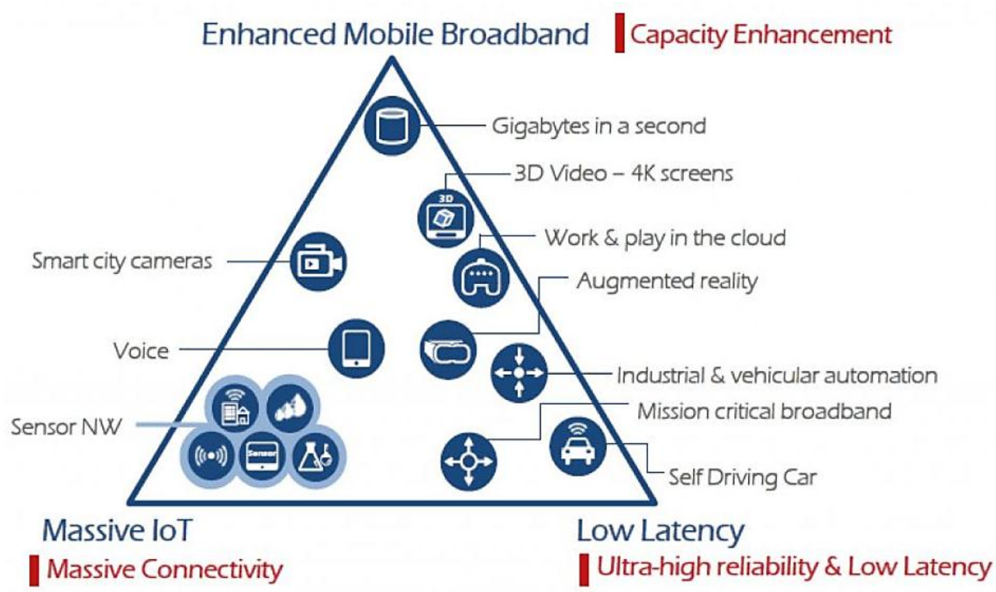


**Figure 1.2**: 5G usage scenarios [4].

## 1.2 Technologies considered in this thesis

In the following, some of the technologies that are relevant for this thesis are discussed:

### 1.2.1 Network Slicing

Network slicing is one of the key capabilities that will enable the required flexibility for supporting the wide range of use cases envisaged for 5G, as it allows multiple logical networks, referred to as network slices, to be created on top of a common shared physical infrastructure. Each network slice can be used to serve a particular service category (e.g., applications with different functional requirements) through the use of specific control plane (CP) and/or user plane (UP) functions [7]. The greater elasticity brought about by network slicing will help to address the cost, efficiency, and flexibility requirements of a wide range of vertical services. Moreover, network slicing will help new services and new requirements to be quickly addressed, according to the needs of the industries [8,9]. Through the network slicing, resources of multi-domain infrastructure networks can be efficiently allocated to multiple network slices according to the requirements of the different use cases [10,11]. The realization of network slicing considers, in the most general case, support for specific features and resources both in the core network part, referred to as Core Network slice, and in the Radio Access Network (RAN) part, referred to as RAN slice, which is the focus of this thesis. An efficient RAN slicing for 5G systems needs to face different challenges in terms of capacity, latency, reliability, and scalability. These challenges are a consequence of the increase in network size and demands for radio resources to be used for the transmissions and of the heterogeneity across applications, networks, and devices.

### 1.2.2 Transmission modes in V2X technology

Vehicle-to-everything (V2X) communication refers to the exchange of information between a vehicle and any entity that may affect the vehicle. It incorporates other more specific types of communication as V2I (vehicle-to-infrastructure), V2N (vehicle-to-network), V2V (vehicle-to-vehicle), and V2P (vehicle-to-pedestrian). V2X is attracting significant attention as it promises to reduce road fatalities, increase traffic efficiency, improve mobility, enable a high-level of vehicle automation, and even reduce environmental impacts and provide additional traveler services. It should be noted that V2X communications have already been standardized by 3GPP, based on LTE Release 14 [12], and described in [1]. LTE- V2X has been designed for specific use cases [9], taking into account services and parameters defined in the first release of ETSI ITS [13] such as safety applications, traffic efficiency, and infotainment services. In current and future releases, 3GPP is working on specifying 5G technologies, and with it, 5G-based V2X. 5G V2X, referred to as enhanced V2X (eV2X) will address more advanced use cases such as cooperative intersection control, lane merging, and platooning, which have more stringent requirements [14].

In LTE-V2X, V2V communications can be performed using two transmission modes, namely cellular mode and Sidelink (SL) (i.e., direct V2V communication) mode. The SL mode refers to the direct V2V communication that allows a User Equipment (UE) at a vehicle to communicate and transmit data directly to other UEs in close proximity over the PC5 interface, bypassing the base station and the network infrastructure. Instead, the cellular mode uses the Uu interface (i.e.

the radio interface between the UE and the base station) with a two-hop transmission via a base station, i.e. involving an uplink (UL) and a downlink (DL) transmission. Integration of the sidelink with cellular networks is a key technology to efficiently support V2V applications and to meet their latency requirements. It is worth mentioning that, although the support for V2X sidelink communications was already standardized in 3GPP in the context of LTE [12], V2X sidelink is not yet included in the current release 15 of 5G New Radio (NR) specifications, but it is subject to study for future release 16 [15].

Direct communication between UEs in cellular networks brings several benefits to both mobile users and network operators such as: (i) Enhancing network performance (i.e., users can experience high data rates, low latency, and reduced energy consumption because of the direct short-range communication and its potentially favorable propagation conditions. (ii) The cellular-coverage range can be extended and the per-area capacity improved without additional infrastructure cost. In fact, cell-edge users, that usually experience poor performance in uplink and downlink transmissions, can communicate directly to nearby terminals or to the base station by means of mobile users acting as relays. (iii) Significant improvement in the spectral efficiency is also foreseen to be achieved by integrating SL communication within the cellular networks. (iv) Allowing spectrum reuse between traditional cellular communications and direct communications which can enhance the spectrum efficiency and allow for a larger number of concurrent transmissions [16,17]. Given the short distance between users, the required transmission power is in general low, which reduces the self-interference level [18]. (v) Because SL communication offers the opportunity for local management of short-distance transmissions, it allows for data offloading from the base station, which alleviates network congestion and traffic management effort at the central nodes [19].

## 1.3 Design challenges

The integration of V2X capabilities in cellular networks poses new technical issues and design problems. In the following, we discuss some of the challenges that V2X faces in future cellular networks.

### 1.3.1 Network Slicing

Network slicing is a fundamental capability for 5G networks to provide a network as a service for different use cases, allowing network operators and to facilitate the cost-effective deployment and operation of multiple logical networks on a shared infrastructure. With network slicing, service providers can deploy their applications and services flexibly and quickly to accommodate the specific requirements of various services. As an emerging technology with a number of advantages, network slicing has raised many issues for the industry and academia alike. The practical realization of such capability still raises numerous technical challenges, both in the Core and RAN parts of the 5G system. Next, we discuss the main challenges arising from implementing network slicing for 5G systems supporting V2X.

- **Isolation among Network Slices:** The demanding requirements of V2X applications, such as ultra-low latency, high-bandwidth, highly-reliable communication, intensive computation and near-real time data processing, raise outstanding challenges and opportunities for 5G systems. As a consequence, dedicated virtual network resources are needed to guarantee the service quality for V2X slice. This requires slices to be highly isolated from other slices. The isolation of V2X from other network slices can be achieved via data plane isolation and control plane isolation. Further, the effective isolation of network slices can ensure that the failure on one slice doesn't affect other slices operation. Hence, the slice isolation mechanism with respect to Quality of Service (QoS) is a major challenge of implementing network slicing.

- **Flexibility and Scalability:** Dynamic slice creation and allocation of network resources management is one of the concerns for future 5G networks to maximize their benefits. To accommodate the maximum number of diversified V2X service requests and achieve high QoS based on V2X requirements, 5G network operators need to allocate network resources rapidly to build network slices. 5G network operators should be able to scale the V2X slice and other slices dynamically according to the varying service load. In addition, management of the network slice should be implemented in an automated fashion to avoid manual efforts and errors. In this context, due to the dynamic characteristics of a V2X service under cellular system, the lifecycle management of the V2X slices is a critical problem and needs to be solved. To solve the problem, we must propose practical resource allocation algorithms with the ability to reconfigure and migrate slice resources dynamically based on V2X requirements.

### 1.3.2 Mode Selection

Mode selection, i.e. selection between the cellular or sidelink mode for a V2V communication, plays an important role in increasing the spectrum efficiency in cellular networks which leads to various benefits i.e., in terms of network capacity utilization, network traffic congestion, the transmission delay, and ensuring a high quality of signal for all users in the network. However, mode selection is crucial and bring some challenges for direct V2V communication between vehicles as an underlay to 5G mobile networks. Following are some design challenges related to the mode selection problem;

- When the mode selection should be triggered and updated in order to achieve the optimization target and improve the performance of the network,
- The information to be considered to decide when it is appropriate to select sidelink or cellular modes,
- Another crucial choice is that the network may have to perform the mode selection very frequently, resulting in a high computational process and significant cost. In addition, the timescale for the mode selection cannot be too coarse because the wireless channel might change rapidly, making the assigned modes inadequate.

To face the challenges mentioned above, efficient mode selection strategies for V2X systems need to be proposed with the ability to ensure the quality of service requirements for all V2X users.

### 1.3.3 Resource Allocation

In LTE - V2X, the transmission time is tightly coupled to the subframe duration, i.e., 1ms. However, if the UE has only a small amount of data to transmit, which can be accommodated in less than 14 OFDM symbols, whenever the packet arrives at the UE for transmission, it must wait until the beginning of the next slot to start transmission. Such a slot-based scheduling is not appropriate for V2X use cases that have latency-critical messages with small payloads. In turn, NR V2X will also support mini-slot scheduling, where UEs can start their transmissions at any of the 14 OFDM symbols and can occupy any number of OFDM symbols within the slot. The use of mini-slots can provide low latency payloads with instant start time without having to wait for the slot boundaries to start. Although short TTI (e.g. 0.25 ms) is useful in reducing latency, it has a cost in terms of higher signaling overhead and therefore lower spectral efficiency [20].

According to the Release 14 of LTE standards, V2V links can be scheduled either by centralized scheduling by base station or distributed scheduling by vehicles [21]. The centralized scheduling provides flexible centralized control over network resources. However, the main challenge of these approaches is that they experience additional delay and more packet drops due to transmission on two hops (i.e., UL and DL). Furthermore, the centralized scheduling needs for the CSI at the base station level which suffers from a trade-off between the large amount of overhead to collect global information (i.e. especially in scenarios where the channels vary rapidly with time) and the imperfect knowledge of the channels states [22] - [25]. In contrast, a decentralized resource allocation scheme has been designed to reduce the latency and ensure high reliability for the vehicular users. However, in a very dense network, decentralized scheduling may experience more packet collisions especially when vehicles are trying to access the same channels at the same time. An efficient radio resource scheduling for V2V systems that benefits from both centralized and distributed scheduling strategies with ability to select the appropriate sizes of TTIs that meets the amount of resources required, as well as the minimum requirements for latency would be useful in order to face different challenges in terms of capacity, latency, reliability, and scalability caused by the increase in network size and demands for radio resources.

## 1.4 Objectives of the thesis

The main objective of this thesis is to study, analyze and propose solutions to address important issues of cellular V2X communications in relation to radio access network slicing, mode selection, and radio resource allocation. This general objective is to be achieved through the following sub-objectives:

1- The proposal and analysis of a novel RAN slicing solution that allocates radio resources to different slices where one of the slices is dedicated to V2X service. The main target of the proposed solution is to maximize the resource utilization while ensuring the availability of resources to fulfill the diverse stringent requirements for V2X service especially in terms of the latency of V2X services, achievable data rate, and outage probability.

2- Investigate the mode selection for cellular V2X communications. In this regard, a novel algorithm for mode selection algorithm is proposed taking into account the quality of the links, available resources, load and the interference situation. This targets to achieve a better network performance by (i) minimizing the consumption of radio resources, (ii) reducing the network traffic congestion, and (iii) ensuring a high quality of signal for all users in the network.

3- Investigate the radio resource allocation in order to improve the V2X service performance in terms of latency, packet success rate and resource utilization. In this regard, a novel solution combining both mode selection and radio resource allocation will be introduced. The proposed solution is composed three novel algorithms for mode selection, a centralized resource allocation, and a distributed resource allocation for V2X networks.

## 1.5 Thesis Structure and Contributions

The remaining part of the thesis is organized as follows.

In chapter 2, we provide some necessary background information on the V2X technology, focusing on the context of V2V technology integrated in cellular systems. This chapter presents the architecture designs for LTE and 5G-based V2X communications starting from the identified novel use cases, technical requirements and limitations. Moreover, we discuss in detail various aspects of network slicing and resource management techniques for V2V-enabled networks, together with the related literature review.

Chapters 3 presents our contribution related to the first research direction that focuses on the RAN slicing management. The chapter starts with an introduction and related works, followed by a thorough explanation of the system model of RAN slicing. This chapter investigates the RAN slicing problem for providing two generic services of 5G, namely enhanced mobile broadband (eMBB) and vehicle-to-everything (V2X). The problem is formulated as an optimization problem to determine the amount of resources assigned to each slice with the target of improving the radio resource utilization while satisfying the specific requirements for each slice. To solve the problem, a novel Radio Access Network RAN slicing strategy based on offline Q-learning followed by a low-complexity heuristic approach is proposed as an enhanced solution to determine the adequate split of resources between the eMBB and vehicle-V2X servces. Furthermore, the performance of the proposed approach is evaluated using extensive simulations to demonstrate its capability to perform an efficient allocation of resources among slices in terms of resource utilization, latency, achievable data rate and outage probability.

Chapters 4 and 5 focus on the mode selection and radio resource allocation management. The chapters start with an introduction and related work, which more specifically illustrates the background and previous work, followed by a thorough explanation of the system model and assumptions in various simulation scenarios.

In Chapter 4, a new a novel mode selection strategy that takes into account the quality of the links between V2V users in sidelink mode and between the base station and vehicles in cellular mode, the available resources, and the network traffic load situation are proposed. Moreover, the impact of Sidelink (SL) communication underlying cellular networks deployment on the performance of the protocol is extensively studied via analysis and simulations. More specifically, a simulation-based analysis is presented to assess the performance of the proposed mode selection strategies to demonstrate its capability to improve the system performance for V2V communication in terms of network capacity, data traffic loads, latency, and probability of outage.

In Chapter 5, we study the resource allocation problem including jointly mode selection and radio resource scheduling with the objective of minimizing the probability of exceeding the maximum delay under the constraint of satisfying the reliability requirement for the V2V communications. To solve the resulting problem, we propose a novel mode selection strategy to decide when it is appropriate to select sidelink mode and use a distributed approach for radio resource allocation or cellular mode and use a centralized radio resource allocation. Furthermore, a novel low-latency resource allocation strategy is proposed for V2V communications using two approaches: distributed radio resource allocation for sidelink mode and centralized radio resource allocation for cellular mode. The performance of the proposed approach is evaluated using extensive simulations to demonstrate its capability to perform an efficient allocation of resources among users in terms of latency, packet success rate and resource utilization.

Finally, chapter 6 discusses the conclusions of the presented work and outlines the potential directions for future investigations.

Figure 1.3 illustrates the main contributions and the structure of the thesis.

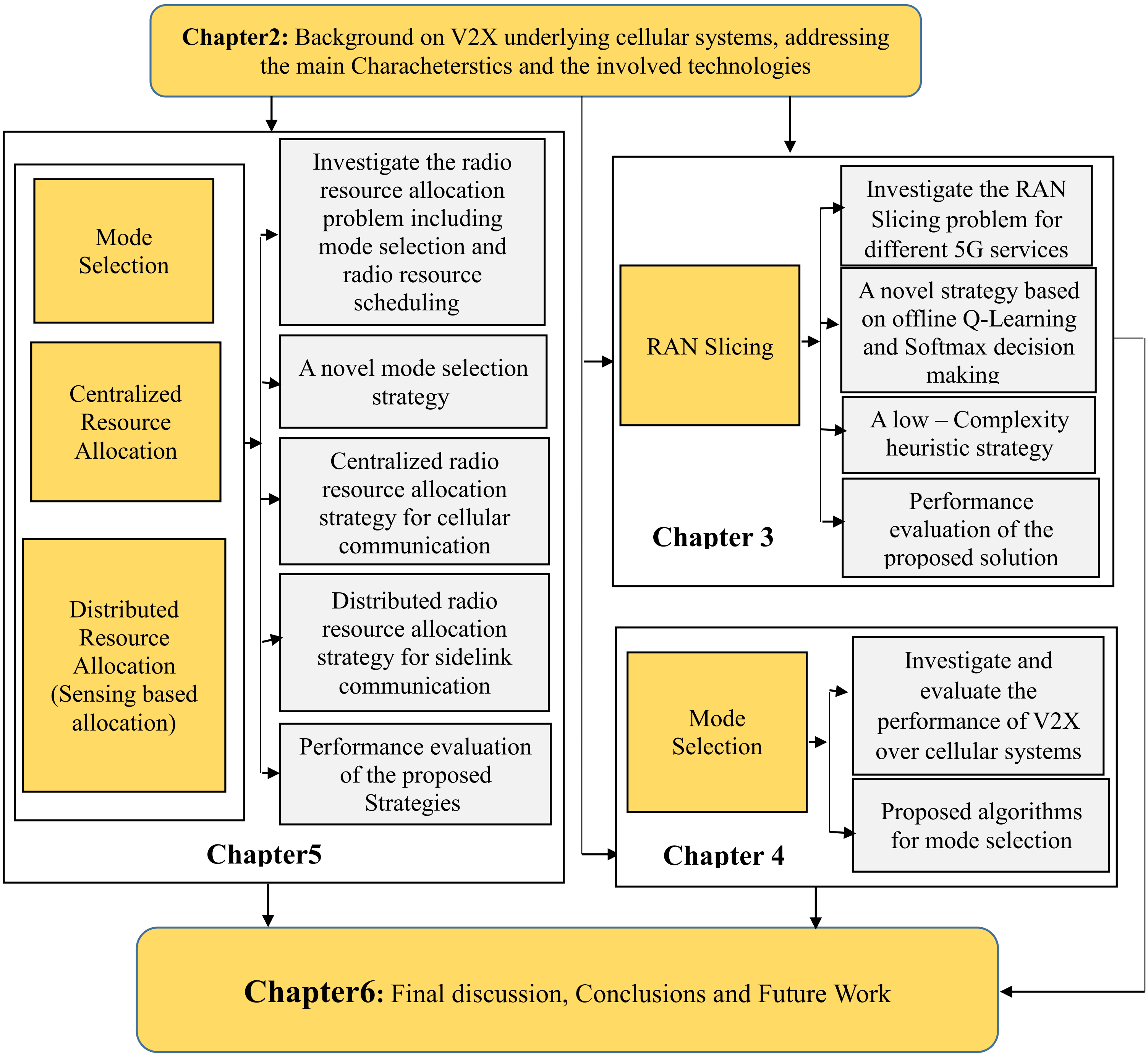


**Figure 1.3:** Structure of the Thesis.

## 1.6 Research Contributions

The following publications have been published in journals and conferences as a result of the work in this thesis.

### Journal Papers

[J1] Haider D. Resin Albonda  Jordi Pérez-Romero , “An Efficient RAN Slicing Strategy for a Heterogeneous Network With eMBB and V2X Services”, IEEE Access, vol. 7, pp. 44771 - 44782, Mar. 2019.
URL: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8676226

[J2] Haider D. Resin Albonda  Jordi Pérez-Romero , “Analysis of RAN Slicing for Cellular V2X and Mobile Broadband Services Based on Reinforcement Learning”, EAI endorsed Transactions on Wireless Spectrum, vol. 4,issue 13, pp. 1- 11, Mar. 2020.
URL: https://eudl.eu/pdf/10.4108/eai.13-7-2018.163841

### Conference Papers

[C1] Haider Daami R. Albonda, J. Pérez-Romero, ``An Efficient Mode Selection for improving Resource Utilization in Sidelink V2X Cellular Networks. IEEE (CAMAD) workshops. Barcelona, Spain, Sep. 2018.
URL: https://ieeexplore.ieee.org/document/8514958

[C2] Haider Daami R. Albonda, J. Pérez-Romero, “Reinforcement Learning-based Radio Access Network Slicing for a 5G System with Support for Cellular V2X”`. International Conference on Cognitive Radio Oriented Wireless Networks (CROWNCOM), Poznan, Poland. 2019.
URL: https://www.springerprofessional.de/en/reinforcement-learning-based-radio-access-network-slicing-for-a-/17056710

[C3] Haider D. Resin Albonda  Jordi Pérez-Romero , “A New Mode Selection and Resource Reuse Strategy for V2X in Future Cellular Networks”, IEEE 91st Vehicular Technology Conference: VTC202, Antwerp, Belgium,  May.2020.

# Chapter 2

# Background in Radio Resource Management for V2X

In this chapter, we first survey the background of V2X communication, focusing on its integration in future cellular networks, use cases, and technical requirements. Then, we discuss various aspects of network slicing and the radio resource management approaches for making the best use of these technologies.

## 2.1 V2X concept

V2X communication enables vehicles to cooperate with other vehicles, devices, and infrastructures. The goal of V2X is to improve road safety, increase traffic efficiency, even reduce environmental impacts and provide traveler information services. Cellular V2X (C-V2X) refers to the provision of V2X services using cellular networks, such as LTE or 5G. The 3GPP initiated standardization activities for C-V2X in LTE in Release 14 [12]. In Release 16 [26], 3GPP further enhances the V2X functionalities to support 5G. As illustrated in Figure 2.1, C-V2X incorporates several components [12], including:

- **Vehicle-to-Vehicle (V2V)**

It covers direct communication between UEs in close proximity. The UE supporting V2V applications transmits application layer information (e.g. about its location, moving speed, and/or attributes as part of the V2V service). V2V is predominantly broadcast-based and includes the exchange of V2V-related application information between distinct UEs directly or via the infrastructure supporting V2X service, (e.g., roadside unit (RSU), and application server) due to the limited direct communication range of V2V.

- **Vehicle-to-Infrastructure (V2I)**

V2I allows vehicles to communicate with external systems such as street lights, buildings, and even cyclists via the RSU. The RSU is a transportation infrastructure entity that could be implemented in a base station (i.e., the eNB in case of LTE or the gNB in case of future 5G NR). The RSU provides several services based on the knowledge of local topology obtained from neighboring vulnerable users, sensors (e.g., cameras, induction loops), and application server. The application server could be deployed within or outside of the cellular network. The UE that supports V2I applications sends application layer information to RSU. In turn, RSU sends this information to a group of UEs supporting V2I applications. The RSU can select the receiving UEs according to the V2I application information in different transmission modes, such as unicast, multicast, and broadcast.

- **Vehicle-to-Network (V2N)**

Vehicle-to-network (V2N) connect vehicles to cellular infrastructure and the cloud so drivers can take advantage of in-vehicle services like traffic updates and media streaming. Some of the most common examples of this technology are vehicles with built-in traffic and navigation functions such as Google Maps, or vehicles that can sync with a smartphone.

- **Vehicle-to-Pedestrian (V2P)**

V2P covers cellular-based communication between UEs supporting V2P applications, where P represents vulnerable road users (Vulnerable Road User) including pedestrians, motorcyclists, bikers, roller skaters, etc. V2P includes the exchange of V2P-related application information between distinct UEs (one for vehicle and the other for pedestrian) directly and/or, due to the limited direct communication range of V2P, the exchange of V2P-related application information between distinct UEs via infrastructure supporting V2X Service, e.g., RSU, application server, etc. V2P application information can be broadcast by a vehicle with UE supporting V2X Service (e.g., warning to pedestrian), and/or by a pedestrian with UE supporting V2X Service (e.g., warning to the vehicle).

### 2.1.1 V2X communication in future cellular network

V2X technology has a great potential of enabling a variety of novel applications for road safety, passenger infotainment, car manufacturer services, and vehicle traffic optimization. There are two communications solutions that have been developed for V2V communication to ensure interoperability in information exchange among vehicles ,namely ad-hoc communications over the IEEE 802.11p standard [3 13] and cellular V2X [12]. In contrast to IEEE 802.11p, C-V2X requires less additional investments in network infrastructure. In addition, the link performance of IEEE 802.11p could also be degraded due to the presence of the hidden node problem [4 1]. The cellular V2X standardized system (i.e, LTE-V2X) can operate in a wide geographical area and meet the requirements of V2X.

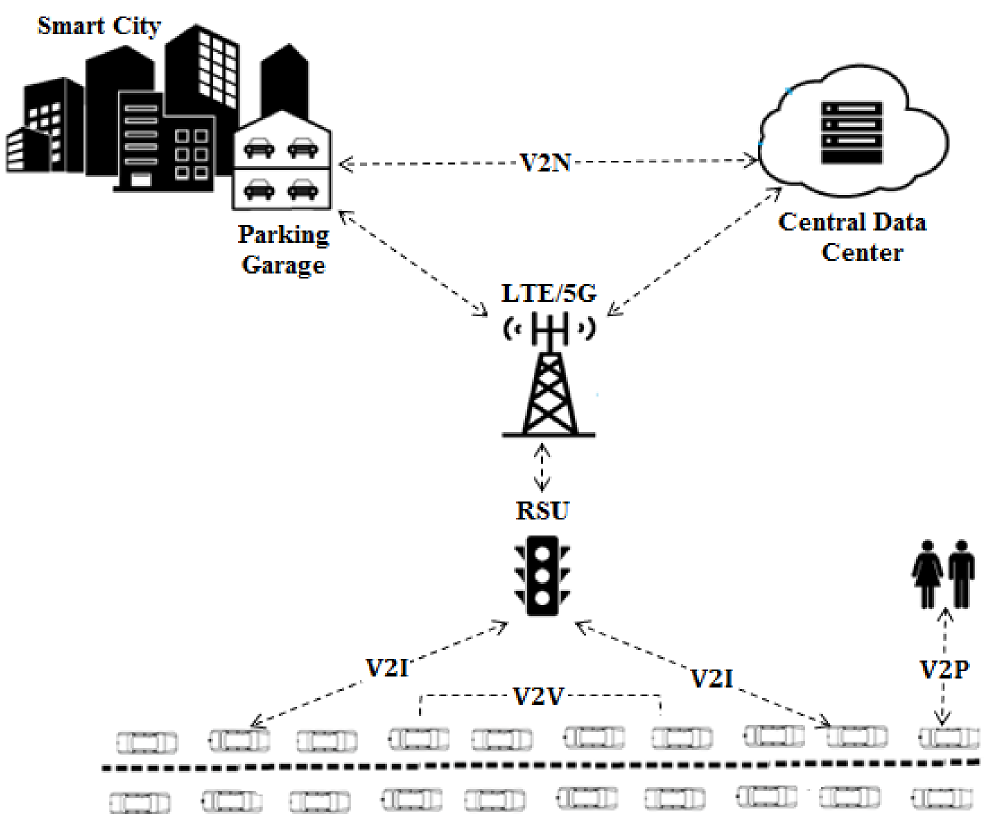


**Figure 2.1:** C-V2X Communications.

LTE has potential to support various V2X services successfully because it has an air interface of high spectral efficiency and able to support different types of communications from one-to-one to one-to-many transmissions, and from conventional uplink and downlink cellular communications to sidelink over-the-air communications.

There are several relevant facts important to be highlighted when comparing IEEE802.11p to LTE-V2X:

- Standardization of LTE-based V2X is being actively conducted to provide the solution for V2X communications that benefit from the global deployment and fast commercialization of LTE systems. Because of the wide deployment of LTE networks, V2I and V2N services can be provided with high data rate, comprehensive quality of service (QoS) support, ubiquitous coverage, and high penetration rate [5 27].
- The existing LTE infrastructure, including eNodeBs and the core networks can be reused with some upgrading in order to provide V2I/N services. A V2I/N service provider can save the cost of deploying new RSUs and connecting them to the network (e.g., the ITS server). With LTE network, a central ITS server could be deployed outside of the LTE network or within the core network (Evolved Packet Core (EPC)) to reduce the latency and to provide centralized control for other entities traffic, road, and service information [28].
- An eNodeB can allocate non-overlapping resources to different UEs in order to prevent resource collision, which is unavoidable in DSRC in a congested area. This eNodeB-based scheduling can be used for both uplink and sidelink transmissions whenever the transmitting UE is inside the network coverage. When eNodeB-based scheduling is not used, the UE can try to avoid resource collision by detecting other UEs' transmission.
- LTE-based V2X supports two air interfaces (cellular interface based on uplink/ downlink and PC5 interface using sidelink, i.e. D2D communication). They will be jointly operated and selected according to the requirement of each V2X service. The joint use of cellular communication and D2D communication will introduce significant operational benefit and efficient utilization of the spectrum, latency, and reliability [12].
- LTE supports one-to-many communications via downlink transmissions from single cell or from multiple cells by using evolved multimedia broadcast multicast services (eMBMS). The one-to-many communication mechanisms in LTE are useful building blocks for V2X

services, especially for safety-related services where the same message needs to be sent to multiple UEs.

- LTE supports the frequency domain multiplexing of multiple UE transmissions in contrast to DSRC, where only one device can transmit at a time in a given channel. As a result, LTE can multiplex more UEs within limited resources without compromising each transmission's coverage, which is especially advantageous when the vehicle density is high.
- LTE can provide better performance when the received signal power is weak. The receiver sensitivity is lower than that of DSRC, which means that LTE UEs can receive weak signals that are not detectable by DSRC receivers. In addition, the use of turbo codes can provide better channel coding gain when compared to the convolutional code used in DSRC.
- IEEE 802.11p has been selected as the technology for V2X communications in some countries in Europe and the United States. The intrinsic characteristics of IEEE 802.11p have confined the technology to support low latency with high reliability [1]. On the other hand, C-V2X has been selected in the U.S., Europe and several regions in Asia. With C-V2X, ubiquitous coverage can be utilized for V2I/V2N services, and interoperability with commercial operators can easily be achieved.
- The link performance of IEEE 802.11p could also be degraded due to the presence of the hidden node problem. However, in LTE-based V2X the hidden node problem can be avoided [1].

Motivated by the aforementioned advantages of the existing cellular network infrastructures, providing better security, longer communication range, and technology evolution path from 4G to 5G and beyond, in this work we consider the cellular systems as a promising technology for V2X communications.

### 2.1.2 Use cases and technical requirements for V2X

In this section we present the advanced use cases that cellular V2X communication will enable to accelerate the development of V2X and we provide precise technical requirements associated with these V2X use cases.

### 2.1.2.1 Service requirements for cellular V2X

This section provides precise technical requirements associated with the V2X use cases that will be later on explained in section 2.1.2.2. The following definitions are utilized to provide specific requirements for V2X services.

- **End-to-end latency (ms):** Maximum tolerable elapsed time from the instant a data packet is generated at the source application to the instant it is received by the destination application. If direct communication (sidelink transport) is used, this is the maximum tolerable air interface latency. Typically, V2X services require communication latency below 100 ms for end-to-end message delivery between the vehicles.

- **Reliability:** Maximum tolerable packet loss rate at the application layer. A packet is considered to be lost if it is not received within the maximum end-to-end latency by the end user. For example, 10-5 means the application tolerates at most 1 in 100,000 packets not being successfully received within the maximum tolerable latency. This is sometimes expressed as a percentage (e.g., 99.999%) elsewhere [29,30].

- **Node mobility** (km/h): Maximum relative speed under which the specified reliability should be achieved.

- **Node density**: Network density (vehicles/$km^2$) is the maximum number of vehicles per unit area under which the specified reliability should be achieved.

- **Offered load:** The highest amount of data traffic on average is expected in urban and highway environments. Urban environments offer a high density of information and thus many objects to be signaled to the car. Highway scenarios can include fast traffic which requires more foresight and thus more information to be signaled to the car. The suburban environment is usually less dense and traffic is comparably slow.

- **Communication range (m):** The maximum distance between source and destination(s) of a radio transmission within which the application should achieve the specified reliability [29]. Assuming urban environments, the cars within the close vicinity are the major interaction partners, and hence a range of 50-100 meters is deemed appropriate. The ranges for the other speeds were chosen such that they scale like the stopping distance of a car (at the corresponding speed).

- **Data rate (Mbit/s):** Minimum required bit rate for the application services. In general, remote processing and all the information captured by the vehicles would require uploading via the communication infrastructure, so the bandwidth requirements can go up to 100 Mbit/s.

#### 2.1.2.2 Use cases for Cellular - V2X

V2X use cases focus on aspects such as safety, traffic efficiency, and infotainment services. Key functional and performance requirements for safety have already been described by the European Telecommunications Standards Institute (ETSI) Intelligent Transport Systems (ITS) [13], the US Department of Transportation, and individual research projects. These use cases are typically used for warning and increasing the environmental awareness based on periodic (e.g., Cooperative Awareness Message – CAM [31]) or event driven (e.g., Decentralized Environmental Notification Message – DENM [32]) broadcast messages, with repetition rate as high as 10 Hz (e.g., emergency vehicle warning) or lower (e.g., road works warning) [12]. V2X use cases have also been identified by 3GPP for LTE-V2X [12], taking into account services and parameters defined in the first release of ETSI ITS [13]. Table 2.1 shows examples of these use-cases. In this group of use cases, the maximum tolerable latency is 100 ms, while the target radio layer message reception reliability is 95%. These use cases assume a single enabling technology, namely cellular based V2X communication. Furthermore, enhanced V2X (eV2X) use cases have been defined by 3GPP as part of Release 16 [14], including more advanced use cases such as cooperative intersection control, lane merging, and platooning, which have more stringent requirements. As shown by table 2.1, the performance requirements are more stringent, with certain use cases requiring ultra-reliable communication links (>99%), with much lower maximum end-to-end latency (1-50 ms), and higher data rate. Table 2.1 shows some of these use-cases specified by [12,14] that require continuous information exchange among vehicles.

### 2.1.3 V2X Architectures

Comprehensive architecture approach for cellular based V2X is important to be developed based on identified novel use cases, technical requirements, and limitations. The 3GPP has already introduced a standardized architecture approach for LTE-V2X in Release 14 [12]. Then, 3GPP has

TABLE 2.1: Performance requirements of examples of V2X use cases.

| LTE-V2X use cases specified by 3GPP TR 22.885 [12] | | | | |
|---|---|---|---|---|
| **Use case** | **Latency requirements** | **Reliability requirements** | **Message size requirements** | **Speed** |
| Emergency vehicle warning | 100ms | 99% | 50-300 Bytes. | 280 km/h |
| Queue Warning | 100 ms | 90% | 50-400 Bytes | 160 km/h |
| Forward Collision Warning | | 99% to 99.99% | 50-300 Bytes | 160 km/h |
| Vulnerable Road User (VRU) Safety | 60 ms | 95% | 1600 Bytes | 100 km/h |
| **5G V2X use cases specified by 3GPP TR 22.886 [14]** | | | | |
| Lane merge (Cooperative maneuver) | 30 ms | 99.99% | 1200-16000 Bytes | 150 km/h |
| Cooperative collision avoidance | 10ms | 99.99% | 2000 Bytes | 160 km/h |
| vehicle platooning | 10ms | 90% | 50-1200 Bytes | 160 km/h |
| Collective perception of environment | 3- 50 ms | 99% to 99.99% | 1600 Bytes | 280 km/h |

made further improvements to the C-V2X architecture in Release 16 [13] to meet the most demanding performance requirements of V2X based 5G networks. In the following, V2X network architectures based on LTE and 5G are discussed.

#### 2.1.3.1 Network Architecture for LTE-V2X

3GPP has adopted Cellular-V2X (C-V2X) as a general term covering all interfaces and release enhancements for V2X communications. C-V2X relies on two air interfaces [33]: (i) the LTE-Uu interface used to support the communication between the vehicles (i.e., the UEs) and the eNBs (i.e., uplink/downlink air interface); and (ii) the PC5 interface, also known as the sidelink, which supports direct communications among V2X UEs. A roadside unit (RSU) supporting V2X communications is attached to the eNB. The LTE-Uu interface supports unicast and broadcast bearers for data transmissions. Specifically, the UEs send unicast messages via the eNB to an application server, which in turn broadcast them via evolved multimedia broadcast multicast service (eMBMS) for all UEs in the relevant geographical area. This mode is suitable for latency – tolerant use cases (e.g., situational awareness, mobility services). V2X communications over the PC5 interface are inherently multicast, typically targeting multiple UEs with a single transmission.

Figure 2.2 illustrates the principles of 3GPP architecture for LTE-Uu and PC5 based V2X communication. The LTE - based V2X architecture introduces the V2X application server as an entity within an operator's core network. The logical functions of this server include requesting BM-SC (Broadcast-Multicast service center) for activating/deactivating/modifying the Multimedia Broadcast Multicast Services (MBMS) bearer, receiving uplink data from the UE over unicast and delivering data to the UE(s) in a target area using Unicast Delivery and/or MBMS Delivery, and providing the V2X service to the UEs.

In the LTE - based V2X architecture, the V2X control function is a logical function that is used for network -related functionalities required for V2X. It's assumed that there is only one logical

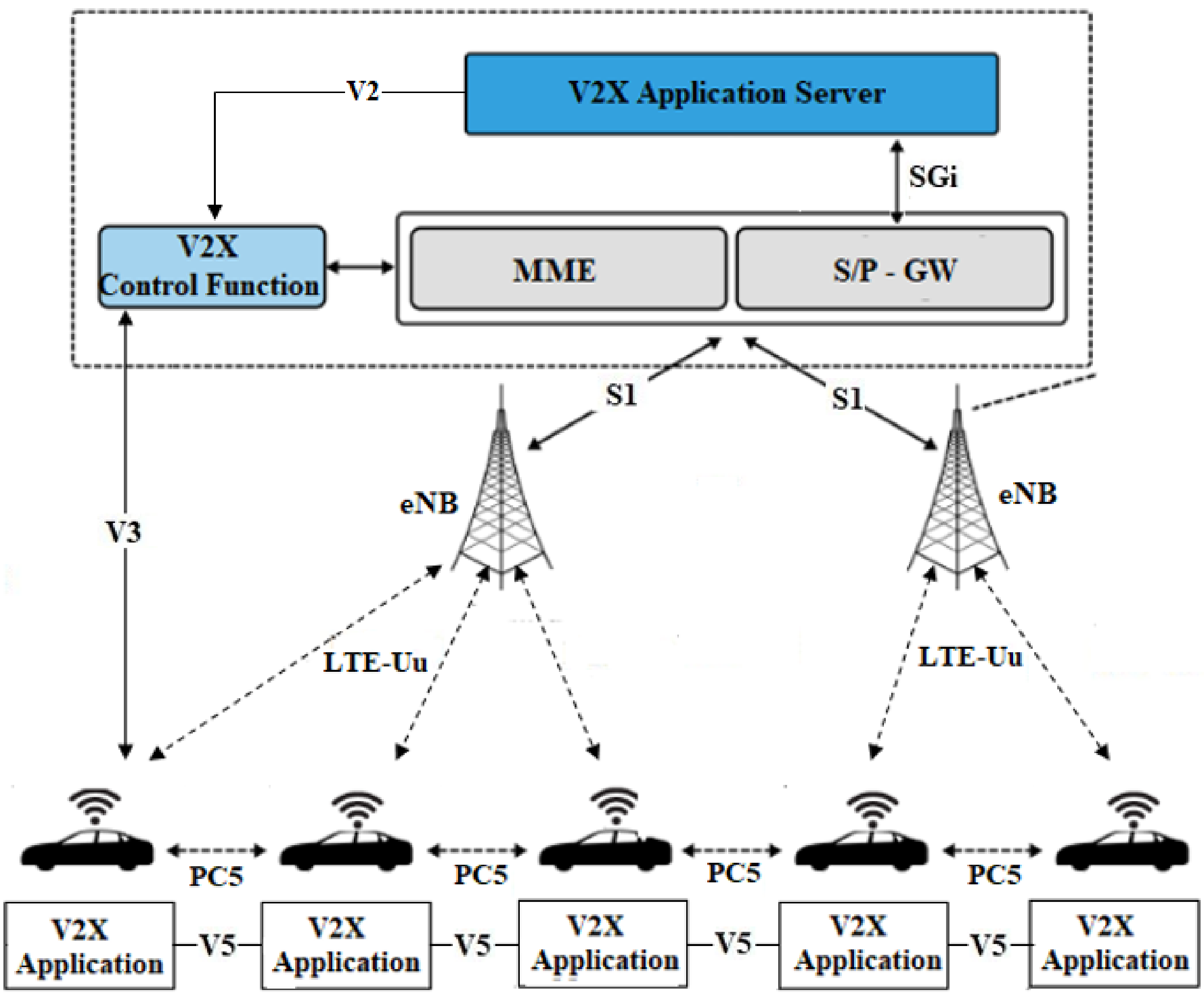


**Figure.2.2:** Architecture for PC5 and LTE-Uu based V2X communication (non-roaming) [33].

V2X control function in each public land mobile network (PLMN) that supports V2X services. The V2X control function is used to configure the UEs with necessary parameters for V2X communications in that PLMN (e.g., the parameters to use PC5 reference point). The V2X control function is also used to configure the UEs with parameters that are needed when the UEs are not served by an LTE network (i.e., V2X UEs out of the cellular coverage and are not able to communicate via the eNB). The V2X control function may also be used to obtain V2X user service descriptions (USDs) from the V2X application server so that the UEs can receive eMBMS-based V2X communication.

#### 2.1.3.2 Network Architecture for 5G-V2X

The 5G system architecture has been designed by 3GPP [26] as a function based architecture which primarily focuses on providing service-based accessibility to the involved entities. In 5G system, architectural enhancements have been specified to support UEs with more advanced V2X services [26]. Figure 2.3 illustrates the V2X architecture that has been specified up to now for 5G [26]. The NG-RAN represents the newly defined radio access network for 5G. NG-RAN provides both NR and LTE radio access. A NG-RAN node is either a gNB (i.e. 5G base station), providing a radio interface (i.e., Uu interface) with the UE based on 5G new radio or ng-eNB, provide a radio interface based on LTE/E-UTRAN. For simplicity, Figure 2.3 depicts only the case of gNBs. The 5G core network includes different network functions such as Policy Control Function (PCF), Access and Mobility Function (AMF), Authentication Server Function (AUSF), Session Management Function (SMF), Application Function (AF), Unified Data Management (UDM), Unified Data Repository (UDR), and User Plane Function (UPF).

Different from LTE-based V2X architecture which required the addition of new functions (i.e., V2X control function), in 5G the V2X functions can be already included in the existing network functions of 5G core. In 5G -based V2X Architecture, PCF is responsible for providing the AMF with necessary parameters in order that the UE can use V2X communication. In more detail, PCF provides the AMF with PC5 QoS parameters and policy parameters to communicate through PC5 and retrieves V2X parameters from UDR. In turn, AMF provides indication about the UE authorization status and the obtained V2X communication information from PCF to the gNB that supports 5G NR via PC5 reference point. V2X connections over the PC5 reference point (i.e., Sidelink) are supported by LTE and / or NR. Although the support for V2X sidelink communications was already standardized in 3GPP in the context of LTE [12], V2X sidelink is not yet included in the current release 15 of 5G New Radio (NR) specifications, but it is subject to study for future release 16 [15].

Furthermore, the UPF is responsible for the enforcement of user plane (UP) policy rules, and in general QoS enforcement. More especially, UPF implements functionalities of packet routing and

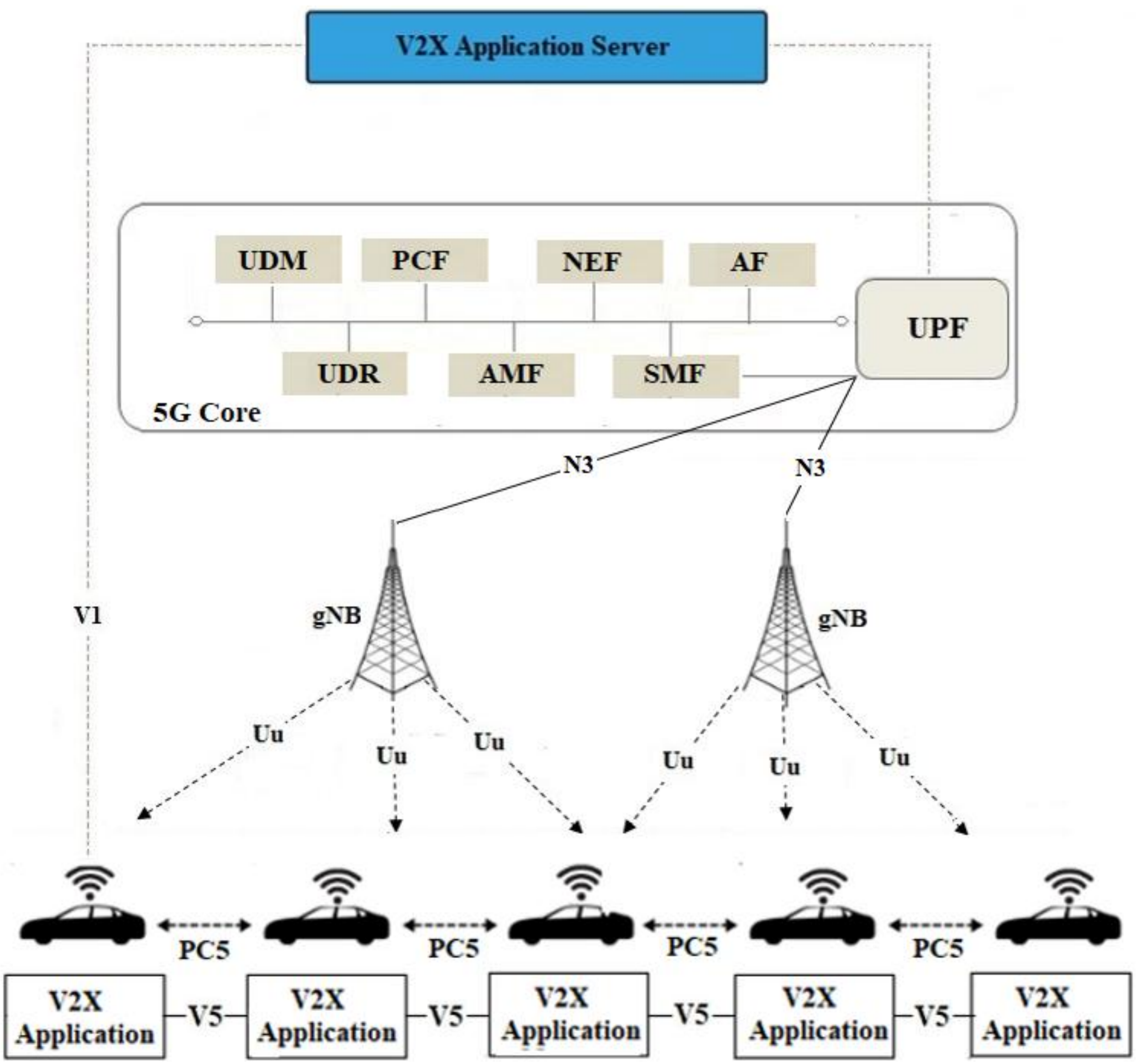


**Figure.2.3:**5G Network Architecture [26].

forwarding. The UPF also acts as an anchor point for intra/inter-RAT mobility, and an external point of interconnection to a data network for Protocol Data Unit (PDU) sessions. The UDM manages information related to the UE and to the subscriber profile. The SMF handles all of the functionalities related to the management of PDU Sessions, such as establishment, modification, and release. Additional functionalities include maintaining tunnels between the NG-RAN and UPF, IP addresses allocation, and management.  The AF is a control plane (CP) function interacting with other NFs of the 5G core (5GC) network and implements procedures such as influence of traffic routing of PDU sessions based on application-specific requirements. AFs may be located within the operator's network and directly connected to the other control plane functions or deployed in a public network. In this case, the AF interacts with the 5GC via NEF, which is responsible for exposing network functionalities and capabilities and handles masking of sensitive internal user information before information exposure.

### 2.1.4 Resource management for V2V-enabled networks

Integration of direct V2V communication in cellular systems creates a need for revisiting the existing radio resource management techniques to make the best use of this technology. From the perspective of this thesis, we consider two main radio resource management approaches for V2V communication in cellular networks: i) Mode selection, that is, deciding if V2V should communicate directly using SL or via gNB (mention also eNB in case of LTE) (i.e., two-hop transmission); ii) resource allocation, that is, assigning the resource blocks (RBs) to the V2V links.

A previously discussed in subsection 1.2.2, a UE can operate in two transmission modes, namely cellular mode and sidelink mode. In this research, the sidelink mode indicates the direct communication between vehicles and cellular mode indicates indirect transmission when two-hop transmission via a base station (or a relay) is established. Mode selection is the problem of choosing whether the users should communicate through sidelink mode or via cellular mode. The optimal mode selection depends on the performance measure to optimize (e.g., sum rate, transmission power, energy consumption), and on the information available when making the decision (e.g., physical distance between users, channel quality of the links, interference level). Note that cellular mode is conceptually the same as a traditional cellular communication.

From the perspective of the radio resource allocation for V2V communication, and according to current LTE standards, V2V services can be supported through two different options, namely mode 3 (i.e., V2V- centralized scheduling) and mode 4 (i.e., V2V-distributed scheduling) to support sidelink communications. These modes are the evolution for V2X of mode 1 (D2D scheduled) and mode 2 (D2D UE-selected) that were introduced in 3GPP Release 12 for Proximity Services (ProSe) supporting public safety communications. In centralized scheduling, the base station (i.e., the eNB in case of LTE or the gNB in case of future 5G NR) selects and schedules the exact resources to be used by the different sidelink transmissions. In distributed scheduling, UEs autonomously select the radio resources for sidelink transmission from a specific pool of resources indicated by the base station. For this purpose, UEs use a sensing-based semi-persistent scheduling (SPS) [34, 35].

In relation to sharing of radio resources, there are basically two approaches for sharing radio resources in sidelink, namely Orthogonal Sharing approach (OS), in which, V2V links can use resources that are orthogonal to those occupied by the other V2V links (i.e., the same RBs are not used by other links), and Non-Orthogonal Sharing approach (NOS), which enables V2V links to use other RBs used by other V2V links. In the OS and NOS approaches, sidelink V2V links get the resource blocks from the resource pools that are allocated for sidelink V2X services and these resources are a subset of available subframes in the UL resources, which are organized based on a periodical subframe pool, as illustrated in Figure 2.4. If users are assigned resources that are orthogonal to those occupied by other users, they cause no interference to each other. However, this is only possible when sufficient amount of resources is available. Therefore, orthogonal resource allocation may result in a lack of the radio resources. This issue may be potentially addressed by leveraging the shared nature of radio resources, which allows the reuse of radio resources when there are no sufficient radio resources for V2V transmissions. The reuse of radio resources may increase spectral efficiency and other metrics such as access delay reduction, but this comes at the cost of introducing interference among the transmissions. Interference makes it challenging to guarantee the reliability requirements of the different services, which calls for the implementation of efficient radio resource management mechanisms.

## 2.2 Network Slicing

5G is expected to provide a great variety of services and applications and to be an end-to-end (E2E) flexible, scalable and demand-oriented system to meet various requirements [35, 36]. Network slicing in 5G is

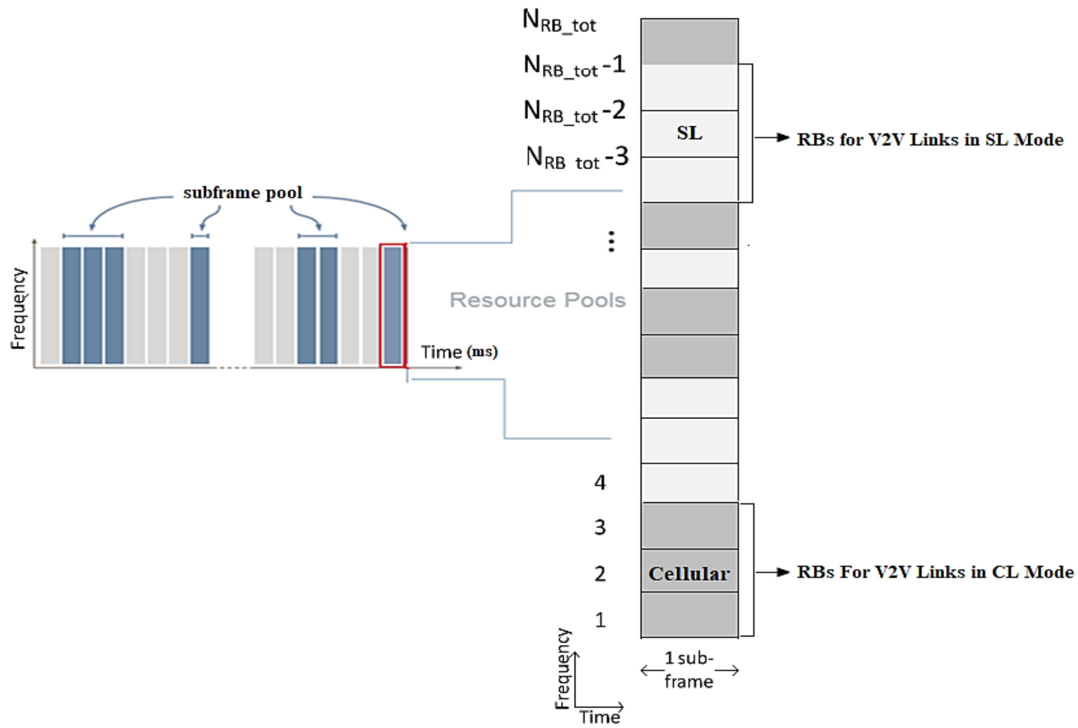


**Figure 2.4:** Organization of radio resources for sidelink transmissions.

considered as a key driving force to achieve this challenging goal and provides a more flexible and efficient implementation to achieve an open 5G network. As explained in subsection 1.2.1, network slicing allows multiple logical networks, referred to as network slices, to be created on top of a common shared physical infrastructure. Each network slice can be used to serve a particular service category in order to meet different user demands. In this respect, 3GPP has completed the normative specifications regarding service and operational requirements to support network slicing [8] and has covered both system architecture aspects [10] and related management and orchestration capabilities [37].

The realization of network slices considers, in the most general case, support for specific features and resources both in the 5G Core Network (5GC) part, which we will refer to as a CN slicing, and in the New Generation RAN (NG-RAN) part, referred to as a RAN slicing. At the Radio Access Network (RAN), the use of network slicing involves the assignment of the radio resources to each slice in accordance with its expected requirements and functionalities (i.e., the demanding requirements of V2X applications, such as ultra-low latency, high-bandwidth, highly-reliable communication, required a dedicated slice highly isolated from other slices to guarantee the service quality for V2X service).

### 2.2.1 Radio Access Network (RAN) Slicing

5G RAN slicing can be implemented through logical abstraction of physical radio resources (such as spectrum) and physical hardware (such as a base station). From the perspective of RAN, the main task of the RAN slice is to achieve the sharing and flexible management of spectrum resources. At the RAN, the use of network slicing involves the assignment of radio resources to each slice in accordance with its expected requirements and functionalities. Therefore, RAN slicing will provide the required design flexibility and will be necessary for any network slicing solution. Specially, the RAN slices can be dynamically configured depending on factors such as traffic type, traffic load and QoS requirement of each. The sharing of radio resources between different slices can be accomplished by scheduling conducted at the MAC layers of the base station [38].

The following design aspects can be considered in RAN slicing [7, 39]:

1) **Resource management:** Radio slices in the RAN can share the radio resources (time, frequency, space) and the corresponding communication hardware (digital baseband processing components, analog radio components) in a dynamic or static manner according to the configuration rules for the network slice. With dynamic resource sharing, each of the slices obtains use of resources based on its demand and priority. The allocation of radio resources among the slices can be done by scheduling. The scheduler allocates radio resources to the slices based on factors such as the quantity of resources requested by the slice, the priority of the service and the overall traffic load.

2) **Slice-specific admission control:** This is needed to meet the initial access requirements of various network slices. RAN configuration rules supply the ability to meet the access needs of a variety of network slices. For example, a network slice serving V2X services must get guaranteed ultra-reliability and low-latency access. With slice-specific admission

control configuration rules, a UE operating in one slice may not be admitted to a gNB if the slice is not active in that gNB. The RAN must be able to control the slice associated with the gNB for V2X service. Specifically, the RAN must be able to consider radio resources and their utilization for the V2X messages of UE to be sent over gNB and provide a means for the MNO to authorize on per subscription basis, the communication range a UE is allowed to use for V2X Service (i.e., provide a means for the MNO to enable or disable the usage of the V2X service of any UE transfer over gNB).

3) **UE awareness on the RAN configurations:** The UEs need to use a service supported by specific network slice have to be aware of the RAN configurations. Especially, the UE needs to receive the RAN configuration prior to access the service. In particular, a UE that supports V2X service must be able to be pre-configurable with V2X slice parameters to be used for the transmission and reception of V2X messages when the UE is not served by gNB (i.e., when UEs out of the coverage of the gNB) which supports V2X Service.

### 2.2.2 Core Network Slicing

5G core network slicing can provide tailored logical networks for different services or verticals in an agile and adaptable fashion [38, 40]. One of the key features in 5G CN is the separation of user plane and control plane capabilities. The key technologies of core network slicing are Software-Defined Networking (SDN) and Network functions virtualization (NFV) [6, 41]. SDN technology can be used to separate the control plane and data plane of the core network so that the control plane and data plane can be deployed independently. The control plane can be centralized, while the data plane can be distributed. For example, the data plane in a network slice for low-latency services can be distributed on the network edge in combination with the Mobile Edge Computing (MEC) technology. MEC technology is a network architecture concept that enables cloud computing capabilities and an IT service environment at the edge of the cellular network and, more in general at the edge of any network (i.e., MEC running applications and performing related processing tasks closer to the cellular customer, network congestion is reduced and applications perform better).

NFV technology can provide the necessary Virtual network functions (VNFs) for the data plane of the core network slices according to the service type. These VNFs can be scaled on-demand as the service changes dynamically.

# Chapter 3

# RAN Slicing Strategy for a Heterogeneous Network with eMBB and V2X Services

## 3.1 Introduction

The traditional mobile communication networks use a one-size-fits-all approach to provide services to mobile devices, regardless the different requirements of the vertical services. Therefore, this design approach cannot support services and use cases with vastly heterogeneous requirements such as high throughput, low latency, high reliability, high mobility, and high security [42]. Thus, it's necessary for the research community to explore new techniques to address the challenges associated with supporting vertical industries. To address the challenges associated with supporting vertical services and use cases, 5G networks need to integrate multiple services with various performance requirements into a single physical network infrastructure, and provide each service with a customized logical network. Network slicing, which allows composing multiple dedicated logical networks with specific functionality running on top of a common infrastructure, is introduced as a solution to cope with this heterogeneity.

Slice-based 5G has the following significant advantages when compared with traditional networks:

- Network slicing can provide logical networks with better performance than one size- fits-all networks.
- A network slice can scale up or down as service requirements and the number of users change.

- Network slices can isolate the network resources of the services in one slice from the others; the configurations among various slices don't affect each other. Therefore, the reliability and security of each slice can be enhanced.
- Finally, through network slicing, resources of a multi-domain infrastructure network can be efficiently allocated to multiple network slices according to the requirements of the different use cases [10, 43].

The realization of network slicing considers, in the most general case, support for specific features and resources both in the core network part, referred to as Core Network slice, and in the Radio Access Network (RAN) part, referred to as RAN slice, which is the focus of this chapter. At the RAN, the use of network slicing involves the assignment of the radio resources to each slice in accordance with its expected requirements and functionalities (e.g., the V2X services that require stringent latency, ultra-reliability). Therefore, RAN slicing will provide the required design flexibility and will be a requisite of any network slicing solution. The need for a dedicated V2X slice has been argued due to the unique V2X use cases' features, by recognizing the poor fitting of reference slices for other traffic types (i.e., eMBB and mMTC). In particular, the potential of the network slicing concept for V2X has been unleashed in [44] and elaborated by 3GPP [45]. By flexibly orchestrating multi-access and edge-dominated 5G network infrastructures, dedicated network slices for V2X safety applications can be isolated over other network traffic. In this regard, this chapter proposes an optimized RAN slicing solution for scenarios with both eMBB and V2X services sharing the same RAN infrastructure. The main focus of the proposed solution is to configure the resource split between the eMBB and V2X slices in order to satisfy their QoS requirements and maximize the system performance for both slices in terms of network metrics such as resource utilization, latency and outage probability.

The proposed solution is considered to allocate resources to users of eMBB services in both UpLink (UL) and DownLink (DL) and to users of V2X services assuming Vehicle-to-Vehicle (V2V) communications. V2V communications can make use of either cellular mode (i.e. using a two-hop transmission via a base station and employing UL/DL of the Uu interface) or SideLink (SL) mode (i.e. nearby vehicles communicate directly over the PC5 interface).

The rest of the chapter is organized as follows. Section 3.2 covers the related works and contributions. In Section 3.3, we present the explanation of the system model of RAN slicing and the problem formulation. Section 3.4 provides the proposed solutions for RAN slicing including an offline reinforcement learning and a low-complexity heuristic algorithm. Section 3.5 presents the performance evaluation followed by the conclusions in Section 3.6.

## 3.2 State of the Art and contributions

There are different works in the literature that have proposed solutions for designing and controlling network slicing [46-51]. A logical architecture for network-slicing-based 5G systems and a scheme for managing mobility between different access networks is proposed in [46]. In turn, the deployment of function decomposition and network slicing as a tool to improve the Evolved Packet Core (EPC) is presented in [47]. They discussed the feasibility of designing a flexible and adaptive mobile core network based on functional decomposition and network slicing concepts. A flexible and programmable Software Defined (SD)-RAN platform is introduced in [48] with a main focus on separating the RAN control and data planes through a southbound API in order to support a flexible control plane for real-time RAN control applications, flexibility and to realize various degrees of coordination among RAN infrastructure entities. Reference [49] focuses on adaptive correction of the forecasted load based on measured deviations, network slicing traffic analysis and prediction per network slice, and admission control decisions for network slice requests. A low complexity heuristic slicing algorithm for joint admission control in virtual wireless networks is proposed in [50]. In [51], a model for orchestrating network slices based on the service requirements and available resources is introduced. They proposed a Markov decision process framework to formulate and determine the optimal policy that manages cross-slice admission control and resource allocation for the 5G networks.

On the other hand, some research studies have been carried out in the context of managing the split of the available radio resources in RAN among different slices to support different services (e.g. eMBB, mMTC, and URLLC) with main focus on the Packet Scheduling (PS) problem in order to improve the network capacity and maximize the sum rates of the whole network using different approaches, such as reinforcement learning [52], auction mechanism [53], and game theory [54]. In [52], a novel radio resource slicing framework for 5G networks with haptic communications is proposed based on virtualization of radio resources. The author adopted a reinforcement learning

(RL) approach for dynamic radio resource slicing in a flexible way, while accounting for the utility requirements of different vertical applications. In this respect, the slicing problem is modeled based on a Markov Decision Process (MDP) and, then, an optimal radio resource slicing strategy based on the application of Q-learning technique is introduced to solve the problem. A network slicing strategy based on a novel auction mechanism is introduced in [53] to decide the selling price of different types of network segments in order to maximize the network revenue and to optimally satisfy the resource requirements.

A network slicing scheme based on game theory for managing the split of the available radio resources in a RAN among different slice types is proposed in [54] to maximize utility of radio resources. Similarly, novel slicing and scheduling schemes are proposed in [55] to meet the requirements such as QoS, fairness, and isolation among different slices. In turn, from an implementation respective, the RB scheduling of different slices has been done by [56]. In order to better support the coexistence of heterogeneous slices with highly diverse QoS requirements and allow heterogeneous utility functions of different slice types in 5G networks, a novel online optimizer based on genetic algorithms has been introduced in [57].

An adaptive algorithm for virtual resource allocation based on Constrained Markov Decision Process is proposed in [58]. Their proposed resource allocation strategy can be dynamically adjusted to allocate power and subcarrier through continuous interaction with the external environment. An online network slicing solution based on multi-armed bandit mathematical model and properties to maximize network slicing multiplexing gains and achieving the accommodation of network slice requests in the system with an aggregated level of demands above the available capacity is proposed in [59]. They provide a novel decisional model addressing the “exploration vs exploitation” dilemma, dubbed as Budgeted Lock-up Multi Armed Problem (BLMAB). The feasibility of the solution is proved through implementation on commercial hardware by considering three network slices such as eMBB for Guaranteed Bit Rate (GBR), eMBB for Best Effort, and Public Safety.

Despite the existence of the abovementioned works in the area of network slicing, none of the above works has considered slicing in vehicular scenarios in which the vehicles can communicate through two operational modes, either sidelink (direct-V2V) via PC5 interface or cellular mode via Uu interface. The consideration of slicing with this type of traffic is indeed one of the novelties of this chapter.

Another novelty of this work relies on the considered methodology for addressing the network slicing problem, making use of a combination of off-line Q-learning with a heuristic method for fine tuning the amount of radio resources allocated to each slice. In this respect, some of the existing solutions in the recent literature have proposed online Q-Learning algorithms for managing the split of the available radio resources among different slices [52]. However, the on-line Q-Learning algorithm needs to keep track of a value table of all possible "actions" that the system can take, and update the values online through an exploration-exploitation process, which may have a chance of making wrong or unevaluated decisions that can degrade actual performance. Instead, to avoid this issue, in our work we focus on offline Q-Learning. This allows exploring the different actions based on a model of the system prior to modifying on-line the actual network configuration. Besides, the Q-learning approach in this work makes use of softmax decision making for selecting the different actions to trade-off between exploration and exploitation, in contrast to the use of ε-greedy selection, which is one of the common techniques used, where ε is the percentage of time the agent takes to select an action randomly rather than taking the action that is most likely to increase its reward given what it knows so far. Although ε-greedy action selection is an effective and popular way to balance exploration and exploitation in reinforcement learning, one of the drawbacks is that when it explores it chooses equally among all actions [60, 61]. This means that the worst-appearing action is likely to be chosen as the best one.

Based on all the above considerations, the key contributions of this chapter are summarized as follows:

- The RAN slicing problem to support an eMBB and a V2X slice on the same RAN infrastructure is formulated as an optimization problem to determine the amount of resources assigned to each slice with the target to improve the radio resource utilization while satisfying the specific requirements for each slice. The model considers uplink, downlink and sidelink (for direct V2V) communications.

- A novel strategy based on offline Q-learning and softmax decision-making is proposed as an enhanced solution to determine the adequate split of resources between the two slices. In this approach, the slice controller continuously interacts with a model of the environment to learn the optimal policy through an immediate reward feedback. The adopted Q-Learning algorithm tracks all possible "actions" (i.e. resource splits) the system can take through an exploration-exploitation process in order to select the appropriate action.

- Starting from the outcome of the off-line Q-learning algorithm, a low-complexity heuristic approach is proposed for fine tuning the resource assignment and achieving further improvements in the use of resources.

- We evaluate the performance of the proposed approach using extensive simulations to demonstrate its capability to perform an efficient allocation of resources among slices in terms of resource utilization, latency, achievable data rate and outage probability. Simulation results show the effectiveness of the proposed schemes under different system parameters.

## 3.3 System Model and Problem Formulation

### 3.3.1 System Model

The considered scenario assumes a cellular Next Generation Radio Access Network (NG-RAN) with a gNodeB (gNB) composed by a single cell. A roadside unit (RSU) supporting V2X communications is attached to the gNB. A set of eMBB cellular users (CUs) numbered as m=1,...,M are distributed randomly around the gNB and a flow of several independent vehicles move along a straight highway, as illustrated in Figure 1. The highway segment is divided into sub-segments (clusters) by sectioning the road into smaller zones according to the length of the road. It is assumed that each vehicle includes a User Equipment (UE) that enables communication with the UEs in the rest of vehicles in the same cluster. Clusters are numbered as j=1,...,C, and the vehicles in the j-th cluster are numbered as i=1,...,V(j). The vehicles in the highway are assumed to enter the cell coverage following a Poisson process with arrival rate $\lambda_v$. The association between clusters and vehicles is managed and maintained by the RSU based on different metrics (e.g. position, direction, speed and link quality) through a periodic exchange of status information.

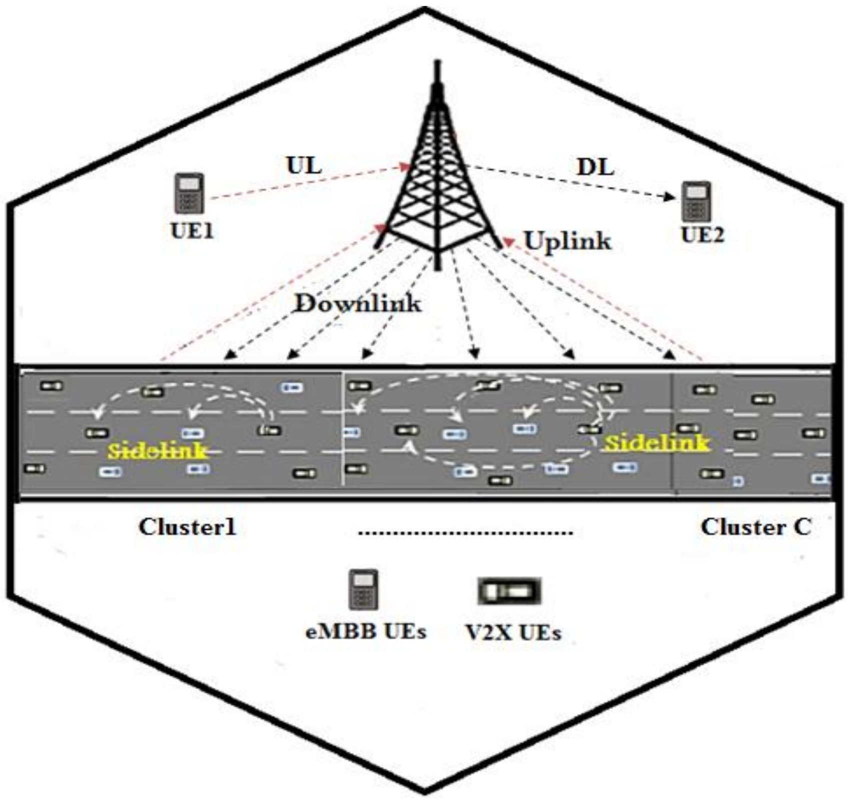


**Figure 3.1**: System model of the cellular network with sidelink V2V.

Regarding the V2X services, this chapter assumes V2V communication between vehicles. It can be performed either in cellular or in sidelink mode. In cellular mode each UE communicates with each other through the Uu interface in a two-hops transmission via the gNB while in sidelink mode, direct V2V communications can be established over the PC5 interface. We assume that, when sidelink transmissions are utilized, every member vehicle can multicast the V2V messages directly to multiple member vehicles of the same cluster $1 \leq i \leq V(j)$ using one-to-many technology. The decision on when to use cellular or sidelink mode will be explained in chapter 4.

### 3.3.2 Network Model for Radio Resource Slicing

In order to jointly support the eMBB and V2X services, and since eMBB requires a large bandwidth to support high- data-rate services, and V2X services are extremely sensitive to latency, the network is logically divided into two network slices, namely RAN_slice_ID=1 for the V2X services and

RAN_slice_ID=2 for the eMBB services. The whole cell bandwidth is organized in Resource Blocks (RBs) of bandwidth B. Let denote as $N_{UL}$ the number of RBs in the UL and $N_{DL}$ the number of RBs in the DL. The RAN slicing process should distribute the UL and DL RBs among the two slices. For this purpose, let denote $\alpha_{s,UL}$ and $\alpha_{s,DL}$ as the fraction of UL and DL resources, respectively, for the RAN_slice_ID=s.

Regarding sidelink communications, and since the support for sidelink has not been yet specified for 5G in current 3GPP release 15, this chapter assumes the same approach as in current LTE-V2X system, in which the SL RBs are part of the total RBs of the UL. For this reason, the slice ratio $\alpha_{s,UL}$ is divided into two slice ratios, $\bar{\alpha}_{s,UL}$, which corresponds to the fraction of UL RBs that are used for uplink transmissions, and $\alpha_{s,SL}$, which corresponds to the fraction of UL RBs used to support sidelink transmissions. The following relationships hold:

$$\sum_{s} \alpha_{s,\mathrm{DL}} = 1 \tag{3.1}$$

$$\sum_{s} \alpha_{s,UL} = \sum_{s} \left( \alpha_{s,SL} + \bar{\alpha}_{s,UL} \right) = 1 \tag{3.2}$$

It is worth mentioning that, since in the considered scenario the sidelink is only used by the V2X slice, it is assumed that $\bar{\alpha}_{2,UL} = \alpha_{2,UL}$ and $\alpha_{2,SL}=0$ for the eMBB slice ($s$=2).

### 3.3.3 V2X Communication Model

Each vehicle is assumed to generate packets randomly with rate $\lambda_g$ packets/s according to a Poisson arrival model. The length of the messages is $S_m$. When the vehicles operate in sidelink mode, the messages are transmitted using the SL resources allocated to the slice. Instead, when the vehicles operate in cellular mode, the messages are transmitted using the UL and DL resources. The average number of required RBs from V2X users of RAN_slice_ID= 1 per Transmission Time Interval (TTI) in UL, DL and SL, denoted respectively as $\Gamma_{1,UL}$, $\Gamma_{1,DL}$, $\Gamma_{1,SL}$, can be estimated as follows:

$$\Gamma_{1,x} = \frac{\sum_{t=1}^{T}\sum_{j=1}^{C}\sum_{i=1}^{V(j)} m(j,i,t)\cdot S_m}{T\cdot SP_{eff,x}\cdot B\cdot T_s} \tag{3.3}$$

where $x$ denotes the type of link, i.e. $x \in$ {UL,DL,SL}, m(j,i,t) is the number of transmitted messages by the vehicles of the $j$-th cluster in the $t$-th TTI and $S_{eff,x}$ is the spectral efficiency in the $x$ link, $T_s$ is the TTI duration and $T$ is the number of TTIs that defines the time window used to compute the average.

### 3.3.4 eMBB Communication Model

Each eMBB user generates sessions requiring a certain guaranteed bit rate. The session generation model follows a Poisson process with rate $\lambda_m$ (sessions/s) and the session duration is exponentially distributed with average $T_e$. These users transmit in the uplink and downlink RBs allocated to the eMBB slice. Then, the average number of required RBs for eMBB users of RAN_slice_ID=2 in UL and DL in order to support a certain bit rate $R_b$ is denoted as $\Gamma_{2,UL}$, $\Gamma_{2,DL}$, respectively, and can be statistically estimated as follows:

$$\Gamma_{2,x} = \frac{\sum_{t=1}^{T}\sum_{m=1}^{M} \rho_x(m,t)}{T} \tag{3.4}$$

where $x$ denotes the type of link, and $\rho_x(m,t)$ is the number of required RBs by the $m$-th user in the link $x$ and in the $t$-th TTI in order to get the required bit rate $R_b$. It is given by $\rho_x(m,t)=R_b/(SP_{eff,x}\cdot B)$. The values $\Gamma_{2,UL}$, $\Gamma_{2,DL}$ are computed within a time window $T$ TTIs Note also that $\Gamma_{2,SL}$=0, since the eMBB slice does not generate sidelink traffic.

### 3.3.5 Problem Formulation for RAN Slicing

The focus of this chapter to determine the optimum slicing ratios $\alpha_{s,UL}$, $\alpha_{s,DL}$ in order to maximize the overall resource utilization under the constraints of satisfying the resource requirements for the users of the two considered slices.

The total utilization of UL resources $U_{UL}$ is given by the aggregate of the required RBs in the UL and SL for each slice, provided that the aggregate of a given slice *s* does not exceed the total amount of resources allocated by the RAN slicing to this slice, i.e. $\alpha_{s,UL} \cdot N_{UL}$. Otherwise, the utilization of slice *s* will be limited to $\alpha_{s,UL} \cdot N_{UL}$ and the slice will experience outage. Based on this, the total utilization $U_{UL}$ is defined as:

$$U_{UL} = \min\left(\Gamma_{s,SL} + \Gamma_{s,UL},\ \alpha_{s,UL} \cdot N_{UL}\right) \tag{3.5}$$

Correspondingly, the optimization problem for the uplink is defined as the maximization of the UL resource utilization subject to ensuring an outage probability lower than a maximum tolerable limit $p_{out}$. This is formally expressed as:

$$\max_{\alpha_{s,UL}} \ U_{UL} \tag{3.6}$$

s.t

$$\Pr\left[\Gamma_{s,SL} + \Gamma_{s,UL} \geq \alpha_{s,UL} \cdot N_{UL}\right] < p_{out} \quad s = 1, 2 \tag{3.6.a}$$

$$\sum_{s} \alpha_{s,UL} = 1 \tag{3.6.b}$$

Following similar considerations like in the uplink, the total resource utilization in the downlink direction $U_{DL}$ is given by

$$U_{DL} = \min\left(\Gamma_{s,DL},\ \alpha_{s,DL} \cdot N_{DL}\right) \tag{3.7}$$

Similar to (3.6), the considered optimization problem of the downlink is formulated as:

$$\max_{\alpha_{s,UL}} \; U_{DL} \tag{3.8}$$

s.t

$$\Pr\left[\, \Gamma_{s,DL} \;\geq\; \alpha_{s,DL} \cdot N_{DL} \,\right] < p_{out} \quad s = 1, 2$$

(3.8.a)

$$\sum_{s} \alpha_{s,DL} = 1 \tag{3.8.b}$$

## 3.4 Proposed Solution for RAN Slicing

The problems in (3.6) and (3.8) with their constraints are nonlinear optimization problems. Such an optimization problem is generally hard to solve. The complexity of solving this problem is high for a network of realistic size with fast varying channel conditions and very tight time-to-decide. For these reasons, we propose the use of an offline reinforcement learning followed by a low-complexity heuristic approach to solve the problem in a more practical way. The general approach is depicted in Figure 3.2. We consider a slicing controller responsible for determining the slicing ratios $\alpha_{s,UL}$, $\alpha_{s,DL}$ for each slice. The operation of the slicing controller is decomposed into two main parts. In the first part, an RL algorithm is responsible for determining some intermediate slicing ratios, denoted as $\beta_{s,UL}$, $\beta_{s,DL}$. Then, the second part is a heuristic algorithm that takes as input the results of the RL algorithm and performs a fine tuning of the slicing ratios in order to determine the final optimized values $\alpha_{s,UL}$, $\alpha_{s,DL}$. This two-step approach allows balancing the trade-off between achieving a fine granularity when setting the slicing ratios and keeping a moderate number of actions in the RL algorithm that facilitates the convergence of the algorithm in a reduced time. A detailed description of the both the RL algorithm and the heuristic algorithm is given in sections 3.4.1, and 3.4.2, respectively.

### 3.4.1 RL-Based Slicing Strategy

It is assumed that two separate RL algorithms are executed for the UL and the DL to determine respectively $\beta_{s,UL}$ and $\beta_{s,DL}$. In the general operation of RL, the optimum solutions are found based on dynamically interacting with the environment based on trying different actions $a_{k,x}$ (i.e. different slicing ratios) selected from a set of possible actions numbered as $k$=1,...,$A_x$, where x $\in${UL, DL}. As a result of the selected action, the RL process gets a reward $R_{TOT,x}(a_{k,x})$ that measures how good or bad the result of the action has been in terms of the desired optimization target. Based on this reward, the RL algorithm adjusts the decision making process to progressively learn the actions that lead to highest reward. The action selection is done by balancing the trade-off between exploitation (i.e. try actions with high reward) and exploration (i.e. try actions that have not been used before in order to learn from them). In case this interaction with the environment was done in an on-line way, i.e. by configuring the slicing ratios on the real network and then measuring the obtained performance, this could lead to serious performance degradation since, during the exploration process, wrong or unevaluated decisions could be made at certain points of time due to the exploration, and affecting all the UEs of a given slice.

To avoid this problem, this chapter considers an off-line RL, in which the slicing controller interacts with a network model (see Figure 3.2) that simulates the behavior of the network and allows testing the performance of the different actions in order to learn the optimum one prior to configuring it in the real network. The network model is based on a characterization of the network in terms of traffic generation, propagation modelling, etc.

The specific RL algorithm considered in this chapter is the Q-learning based on softmax decision making [60], which enables an exploration-exploitation traversing all possible actions in long-term. In turn, the reward should be defined in accordance with the optimization problem, which in this work intends to maximize the resource utilization subject to the outage probability constraint. The details about the reward function and the detailed operation of the Q-learning algorithm are presented in the following.

#### 3.4.1.1 Reward Computation

The reward function should reflect the ability of the taken action to fulfill the targets of the optimization problems (3.6) and (3.8). Based on this, and for a given action $a_{k,x}$ with associated slicing ratios $\beta_{s,x}(k)$ the reward is computed as function of the normalized resource utilization $\Psi_{s,x}(a_{k,x})$ of slice $s$ in link $x\in${UL,DL} defined as the ratio of used resources to the total allocated resources by the corresponding action. For the case of the V2X slice ($s$=1), it is defined as:

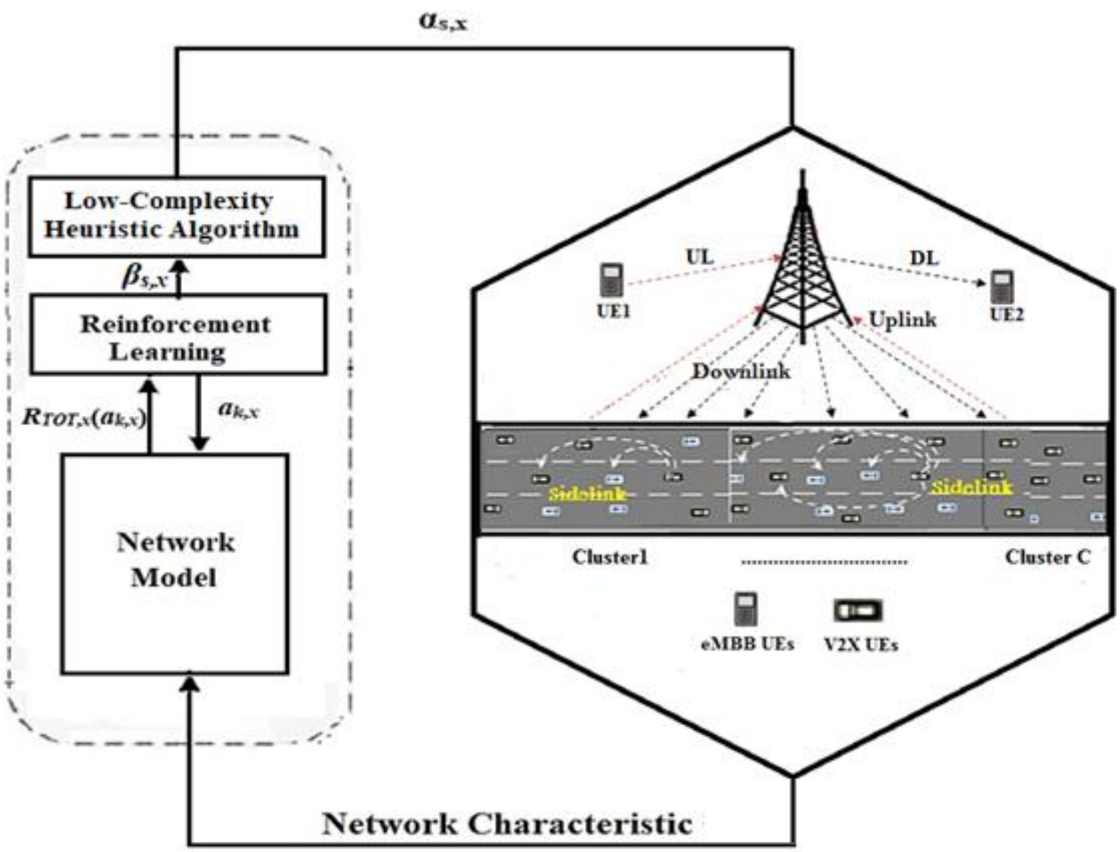


**Figure 3.2**: RAN Slicing Strategy.

$$\Psi_{1,UL}\left(a_{k,\mathrm{UL}}\right) = \frac{\Gamma_{1,UL} + \Gamma_{1,SL}}{\beta_{1,UL}\left(k\right)\cdot N_{UL}} \tag{3.9}$$

$$\Psi_{1,DL}\left(a_{k,DL}\right) = \frac{\Gamma_{1,DL}}{\beta_{1,DL}\left(k\right)\cdot N_{DL}} \tag{3.10}$$

In turn, for the case of eMBB slice ($s$=2), it is defined as:

$$\Psi_{2,UL}\left(a_{k,\mathrm{UL}}\right) = \frac{\Gamma_{2,UL}}{\beta_{2,UL}\left(k\right)\cdot N_{UL}} \tag{3.11}$$

$$\Psi_{2,DL}\left(a_{k,\mathrm{DL}}\right)=\frac{\Gamma_{2,DL}}{\beta_{2,DL}\left(k\right)N_{DL}} \tag{3.12}$$

Based on these expressions, the reward $R_{s,x}(a_{k,x})$ for the slice $s$ in link $x\in\{\mathrm{UL,DL}\}$ as a result of action $a_{k,x}$ is defined as

$$R_{s,x}\left(a_{k,x}\right)=\begin{cases} e^{\Psi_{s,x}\left(a_{k,x}\right)} & \Psi_{s,x}\left(a_{k,x}\right)\leq 1 \\ 1/\Psi_{s,x}\left(a_{k,x}\right) & \Psi_{s,x}\left(a_{k,x}\right)\geq 1 \end{cases} \tag{3.13}$$

In (3.13), whenever $\Psi_{s,x}(a_{k,x})$ is a value between 0 and 1, the reward function will increase exponentially to its peak at $\Psi_{s,x}(a_{k,x})$ =1. Therefore, the actions that lead to higher value of $\Psi_{s,x}(a_{k,x})$ (i.e. higher utilization) provide larger rewards and therefore this allows approaching the optimization target of (3.6) and (3.8). In contrast, if the value of $\Psi_{s,x}(a_{k,x}) > 1$, it means that the slice $s$ will be in outage and thus the reward decreases to take into consideration constraints (3.6a) and (3.8a). Consequently, the formulation of the reward function per slice in (3.13) takes into account the constraints of the optimization problem. In addition, since the total reward has to account for the effect of the action on all the considered slices $s$=1,..,$S$, it is defined in general as the geometric mean of the per-slice rewards, that is:

$$R_{TOT,x}\left(a_{k,x}\right)=\left(\prod_{s=1}^{S}R_{s,x}\left(a_{k,x}\right)\right)^{\frac{1}{S}} \tag{3.14}$$

### 3.4.1.2 Computation of the Q-values and probability selection criterion

The ultimate target of the Q-learning scheme at the slicing controller is to find the action (i.e. the slicing ratios for a given link $x\in\{\mathrm{UL,DL}\}$) that maximizes the expected long-term reward to each slice. To achieve this, the Q-learning interacts with the network model over discrete time-steps of fixed duration and estimates the reward of the chosen action. Based on the reward, the slice controller keeps a record of its experience when taking an action $a_{k.x}$ and stores the action-value function (also referred to as the Q-value) in $Q_x(a_{k,x})$. Every time step, the $Q_{UL}(a_{k,UL})$ and $Q_{DL}(a_{k,DL})$ values are updated following a single-state Q-learning approach with a null discount rate [20] as follows:

$$Q_x(a_{k,x}) \leftarrow (1-\alpha)\, Q_x(a_{k,x}) + \alpha \,.\, R_{TOT,x}(a_{k,x}) \tag{3.15}$$

where $\alpha \in (0, 1)$ is the learning rate, and $R_{TOT,x}(a_{k,x})$ is the total reward accounting for both V2X and eMBB slices after executing an action $a_{k,x}$. At initialization, i.e. when action $a_{k,x}$ has never been used in the past, $Q_x(a_{k,x})$ is initialized to an arbitrary value (e.g. 0 in this work).

### 3.4.1.3 Selection Criterion

The selection of the different actions based on the $Q_x(a_{k,x})$ is made based on the softmax policy [20], in which the different actions are chosen probabilistically. Specifically, the probability $P(a_{k,x})$ of selecting action $a_{k,x}$, $k=1,\dots,A_x$, is defined as

$$P(a_{k,x}) = \frac{e^{Q_x(a_{k,x})/\tau}}{\sum_{j=1}^{A_x} e^{Q_x(a_{j,x})/\tau}} \tag{3.16}$$

where $\tau$ is a positive integer called temperature parameter that controls the selection probability. With a high value of $\tau$, the action probabilities become nearly equal. However, a low value of $\tau$ causes a greater difference in selection probabilities for actions with different Q-values. Softmax decision making allows an efficient trade-off between exploration and exploitation, i.e. selecting with high probability those actions that have yield high reward, but also keeping a certain probability of exploring new actions, which can yield better decisions in the future.

The pseudo-code of the proposed RL-based RAN slicing algorithm is summarized in Algorithm 3.1. Once the offline RL algorithm has converged, i.e. the selection probability of one of the actions is

**Algorithm 3.1** : RAN slicing algorithm based on RL

1. **Inputs:**
 $N_{UL}$, $N_{DL}$: Number of RBs in UL and DL.
 $S$: number of slices,
 Set of actions $a_{k,x}$ for link $x\in\{UL,DL\}$
2. **Initialization of Learning**:
 $t \leftarrow 0$ , $Q_x(a_{k,x})=0$, $k=1,\ldots,A_x$, $x\in\{UL,DL\}$
3. **Iteration**
4. **While** learning period is active do
5. **for** each link $x\in\{UL,DL\}$
6. Apply softmax and compute $P_x(a_{k,x})$ for each action $a_{k,x}$ according to (3.16);
7. Generate an uniformly distributed random number $u \in \{0,1\}$
8. Select an action $a_{k,x}$ based on u and probabilities $P_x(a_{k,x})$
9. Apply the selected action to the network and evaluate $\Psi_{s,x}(a_{k,x})$ based on (3.9)- (3.12).
10. **If** $\Psi_{s,x}(a_{k,x})\leq 1$ then
11. $R_{s,x}\left(a_{k,x}\right)=e^{\Psi_{s,x}\left(a_{k,x}\right)}$
12 **else**
13. $R_{s,x}\left(a_{k,x}\right)=1/\Psi_{s,x}\left(a_{k,x}\right)$
14. **End**
15. **Compute** $R_{TOT,x}(a_{k,x})$ based on equation (3.14)
16. **Update** $Q_x(a_{k,x})$ based on equation (3.15)
18. **End**
19. **End**

higher than 99.99%, the slicing ratios $\beta_{s,x}$ associated to this action are passed to the low-complexity heuristic algorithm, as illustrated in Figure 3.2. Let's denote this selected action as $a_{x\_sel}$.q

### 3.4.2 Low-Complexity Heuristic Strategy

In this section, we propose a heuristic scheme for fine tuning the initial slicing ratios $\beta_{s,UL}$, $\beta_{s,DL}$ chosen by the RL based on the resource requirements. The idea of this fine tuning is that, based on the actual RB demands of each slice and the slicing ratios $\beta_{s,UL}$, $\beta_{s,DL}$ the algorithm assesses if one of the two slices $s$ has more resources than actually required in the link $x\in\{UL,DL\}$, i.e. $\Psi_{s,x}(a_{x\_sel})$ <1, and at the same time the other slice $s'$ has less resources than required, i.e. $\Psi_{s',x}(a_{x\_sel})$>1. If this

**Algorithm 3.2** : Low-complexity Heuristic solution

**Input:**
Slicing ratios selected by the RL: $\beta_{s,x}$, s=1,2. $x \in \{UL, DL\}$ .
**For each** Slice s = {1,2}
**Compute** the required RBs ( $\Gamma_{s,x}$) from equation (3.3) and (3.4)
**Compute** $\Psi_{s,x}( a_{x\_sel} )$ based on (3.9)-(3.12).
**End;**
**If** $\Psi_{1,x}(a_{x\_sel}) < 1$ and $\Psi_{2,x}(a_{x\_sel}) > 1$
$\Delta C_{2,x} = \left(1 - \Psi_{2,x}\left(a_{x\_sel}\right)\right).\ \omega$
$\alpha_{2,x} = \beta_{2,x} - \Delta C_{2,x}$
$\alpha_{1,x} = \beta_{1,x} + \Delta C_{2,x}$
**else if** $\Psi_{1,x}(a_{x\_sel}) > 1$ and $\Psi_{2,x}(a_{x\_sel}) < 1$
$\Delta C_{1,x} = \left(1 - \Psi_{1,x}\left(a_{x\_sel}\right)\right).\ \omega$
$\alpha_{1,x} = \beta_{1,x} - \Delta C_{1,x}$
$\alpha_{2,x} = \beta_{2,x} + \Delta C_{1,x}$
**else**
$\alpha_{1,x} = \beta_{1,x}$
$\alpha_{2,x} = \beta_{2,x}$
**End**

is the case, the slice *s* leaves some extra capacity $\Delta C_{s,x}$ that can be transferred to the other slice *s'*. Specifically, the extra capacity is defined as:

$$\Delta C_{s,x} = \left(1 - \Psi_{s,x}\left(a_{x\_sel}\right)\right).\ \omega \tag{3.17}$$

where the configuration parameter ω is a scalar in the range [0,1] used to leave some margin capacity to cope with the variations of the RBs consumption. Based on this, the slicing ratio for the slice *s* will be decreased in $\Delta C_{s,x}$, i.e. $\alpha_{s,x} = \beta_{s,x} - \Delta C_{s,x}$, while the slicing ratio for the other slice s' will be increased in $\Delta C_{s,x}$, i.e. $\alpha_{s',x} = \beta_{s',x} + \Delta C_{s,x}$. The detailed steps of the proposed heuristic algorithm are shown in Algorithm 3.2.

## 3.5 Performance Evaluation

In this section, we evaluate the performance of the proposed RAN slicing solution through system level simulations performed in MATLAB.

### 3.5.1 Simulation Setup

Our simulation model is based on a single-cell hexagonal layout configured with a gNB. The gNB supports a cell with a channel organized in 200 RBs composed by 12 subcarriers with subcarrier separation $\Delta f$=30 kHz, which corresponds to one of the 5G NR numerologies defined in [62]. The model considers vehicular UEs communicating through cellular mode (uplink / downlink) and via sidelink (direct V2V) and using RAN_slice_ID=1 and eMBB UEs operating in cellular mode (uplink / downlink) and using RAN_slice_ID=2 based on the assumptions described in section 3.3.
The eMBB UEs are randomly distributed in the cell, while the V2X users move along a 3-lane highway. All relevant simulation parameters are summarized in Table 3.1.
As for the RAN slicing approach, the actions of the RL are defined such that action $a_{k,x}$ corresponds to $\beta_{1,x}(k)$= 0.05·$k$ and $\beta_{2,x}(k)$=(1- 0.05·$k$) for $k$=1,...,20, $x \in$ { UL, DL}. The uplink slicing ratio for slice 1 $\alpha_{1,UL}$ obtained as a result of the proposed RAN slicing algorithm is split into two ratios ($\bar{\alpha}_{s,UL}$=0.35·$\alpha_{1,UL}$ for V2X users in cellular mode and $\alpha_{1,SL}$ =0.65·$\alpha_{1,UL}$ for V2X in sidelink mode).
The presented evaluation results intend to assess and illustrate the performance of the proposed RAN slicing solution in terms of RB utilization, throughput, outage probability and latency. As a reference for comparison, we assume a simpler RAN slicing strategy denoted as a proportional slicing strategy (PSS), in which the ratio of RBs for each slice is proportional to its total traffic rate (in Mb/s) and a reference scheme in which a fixed slicing ratio is allocated to each service for uplink (UL+SL) and downlink (70 % of PRBs for slice ID=1 and 30 % of PRBs for slice ID=2), denoted as a fixed slicing strategy (FSS). In addition, in order to see the impact of the two steps in the proposed RAN slicing approach, the results analyze the performance for the case that only the RL is considered, denoted as QL slicing strategy (QL-SS), and the performance for the case that both the RL and the heuristic algorithm are considered, denoted as QL followed by the low-complexity heuristic slicing strategy (QLH-SS).

TABLE 3.1: Simulation parameters

| Parameter | Values |
|---|---|
| **General parameters** | |
| Cell radius | 500m |
| Number of RBs per cell | $N_{UL}$=$N_{DL}$=200 RBs |
| Frequency | 2.6 GHz |
| Path loss model | The path loss and the LOS probability for cellular mode are modeled as in [63]. In sidelink mode, all V2V links are modeled based on freeway case (WINNER+B1) with hexagonal layout [ITU-R] [64]. |
| Spectral efficiency model to map SINR. | Model in section A.1 of [65]. The maximum spectral efficiency is 8.8 b/s/Hz. |
| Shadowing standard deviation | 3 dB in LOS and 4 dB in NLOS. |
| height of the gNB | 10m |
| Base station antenna gain | 5 dB |
| TTI duration ( $F_d$) | 1ms |
| Time window T | 3s |
| **V2X parameters** | |
| Length of the highway | 1Km |
| Number of lanes | 3 in one direction |
| Lane width | 4 m |
| Number of clusters | 4 |
| Size of cluster | 250m |
| Vehicular UE height | 1.5m |
| Vehicle speed | 80 Km/h |
| Vehicle arrival rate $\lambda_v$ | 1 UE/s |
| Packet arrival rate $\lambda_g$ | 1 packets/s |
| Message size ($S_m$) | 300 bytes |
| **eMBB parameters** | |
| UE height | 1.5m |
| Average session generation rate $\lambda_m$ | Varied from 0.2 to 1.2 sessions/s |
| $R_b$ | 1 Mb/s |

| | |
|---|---|
| Average session duration | 120 s |
| **RAN slicing algorithm parameters** | |
| Learning rate α | 0.1 |
| Ω | {0.25, 0.55, 0.85} |
| Temperature parameter τ | 0.1 |
| Actions of the RL algorithm | 20 actions, $k$=1,...,20<br>Slice 1: $\beta_{1,x}(k)=0.05\cdot k$<br>Slice 2: $\beta_{2,x}(k)=(1- 0.05\cdot k)$ |

## 3.5.2 Impact of the parameter ω on the Network Performance Metrics

In this subsection, we study the impact of the parameter ω for the heuristic algorithm, on the network performance metrics i.e., in terms of the obtained RB utilization and throughput. In Figure 3.3 and Figure.3.4, the QLH-SS, considering different values of ω is compared with the QL-SS.

Figure 3.3 plots the obtained RB utilization for UL, as a function of the eMBB session arrival rate ($\lambda_m$) when the number of actions is 20. Since SL and UL make use of the same set of RBs, the results included in Figure 3.3 refer to the total utilization by both links for V2X and eMBB slices. From the presented results, we notice that, when increasing the value of ω, the system provides more resources and therefore leads to better utilization, as it is observed when comparing the results for ω equal to 0.85 against the results for other values of ω.

Regarding the quantitative comparison between strategies, the figure reflects that, for the case of QLH-SS- ω = 0.85, the system utilizes around 94 % of radio resources in uplink when the eMBB session arrival rate is 0.8 sessions/s. In case of QLH-SS- ω =0.55, the system utilizes around 82 % of radio resources in uplink (i.e. QLH-SS with ω = 0.85 achieves a relative gain of 14 %). while for the QLH-SS- ω = 0.25, the system utilization is about 70% (i.e. QLH-SS with ω = 0.85 achieves a relative gain of 34 %). In contrast, in case of the QL-SS, the system utilizes around 60 % of radio resources in uplink (i.e. QLH-SS with ω = 0.85 achieves a relative gain of 56 %).

Figure.3.4 presents the aggregate throughput delivered in Mbits/sec for both eMBB and V2X slices in the sidelink and uplink. From Figure.3.4, we can observe that the QLH-SS with all values of ω outperforms the QL-SS. Specifically, QLH-SS with ω = 0.85 achieves a throughput of 123 Mb/s when the eMBB session arrival rate is 0.8 sessions/s. In turn, the system achieves a throughput of 111 Mb/s in uplink in case of QLH-SS with ω = 0.55 (i.e. QLH-SS with ω = 0.85 achieves a relative gain of 10 %). and 100 Mb/s in case of QLH-SS with ω = 0.25 (i.e. QLH-SS with ω = 0.85 achieves a relative gain of 23 %). The QL-SS achieves a throughput of 90 Mb (i.e. QLH-SS with ω = 0.85 achieves a relative gain of 36% with respect to QL-SS). The reason for this behavior is that, as the number of eMBB sessions increases, requiring more radio resources, the QLH-SS ensures more RBs and achieves higher radio resource utilization compared to the QL-SS. Therefore, these RBs can be used to transmit more data.

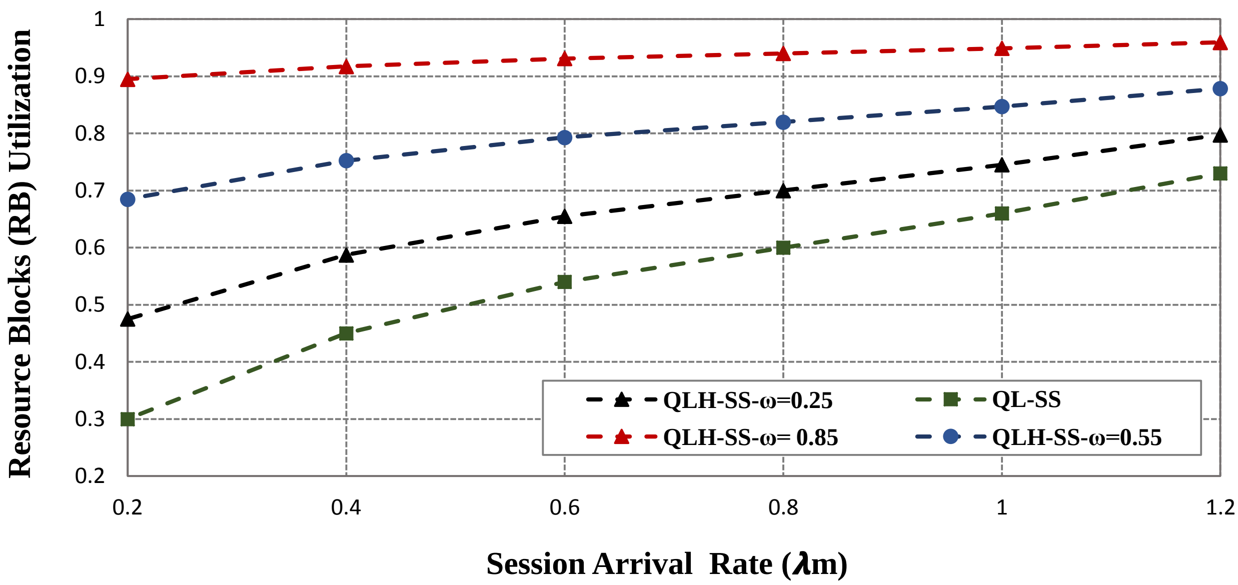


**Figure 3.3:** Uplink RB utilization as a function of the eMBB session generation rate $\lambda_m$ (sessions/s).

### 3.5.3 Comparison of the performance of QLH-SS against reference schemes

#### 3.5.3.1 Performance in terms of RBs Utilization

In this subsection, the performance of QLH-SS with ω = 0.85 in terms of RBs utilization is compared with the QL-SS, PSS, and FSS schemes. Figure 3.5 and Figure 3.6 present the obtained

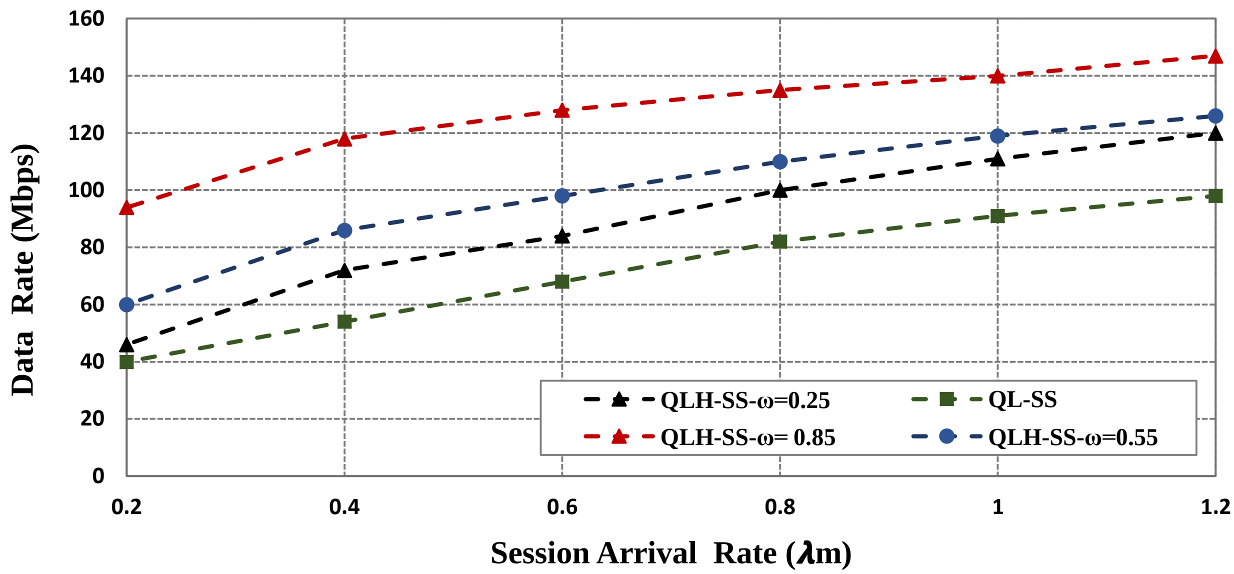


**Figure 3.4**: Aggregated throughput experienced by both slices in uplink as a function of the eMBB session generation rate $\lambda_m$ (sessions/s).

RB utilization, i.e. the number of used RBs normalized to the number of total available RBs, in UL and DL, respectively, as a function of the eMBB session arrival rate ($\lambda_m$). The results included in Figure 3.5 refer to the total utilisation by both links for V2X and eMBB slices. From the presented results, we notice that the QLH-SS and QL-SS approaches maintain high resource utilization compared to PSS, and FSS approaches in different load scenarios. This is due to the RL-based slicing strategy that inherently tackles slice dynamics by selecting the most appropriate action. Further improvements are obtained by the QLH-SS approach by checking the unused capacity left by each slice after selecting an action and use it to serve more traffic load in the other slice. From Figure 3.5 and Fig. 3.6, it's clearly observed that, as the arrival rate of requests increases, the RB utilization of the system increases gradually. For the QL-SS, when the session generation rate of eMBB traffic is 1.2 sessions/s, the system utilizes around 78 % and 74 % of radio resources in uplink and downlink, respectively. Further improvements are obtained by the off-line RL followed by the QLH-SS as the system utilizes up to 91 % of radio resources in uplink and 96 % in downlink. This is due to the benefit from the unused capacity left from each slice. For PSS approach, the utilization is only about 69 % in uplink (i.e. QLH-SS achieves a relative gain of 32 %) and 76 % in downlink (i.e. QLH-SS achieves a relative gain of 20 %). In case of FSS, the utilization is only about 50 % of radio resources

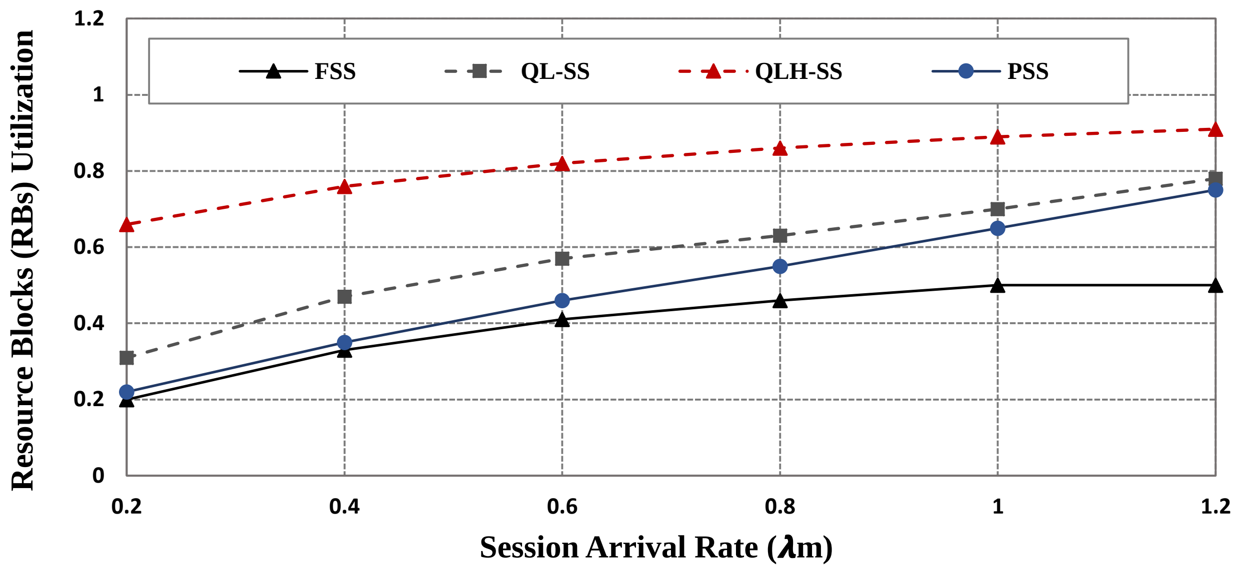


**Figure 3.5**: Uplink RB utilization as a function of the eMBB session generation rate $\lambda_m$ (sessions/s).

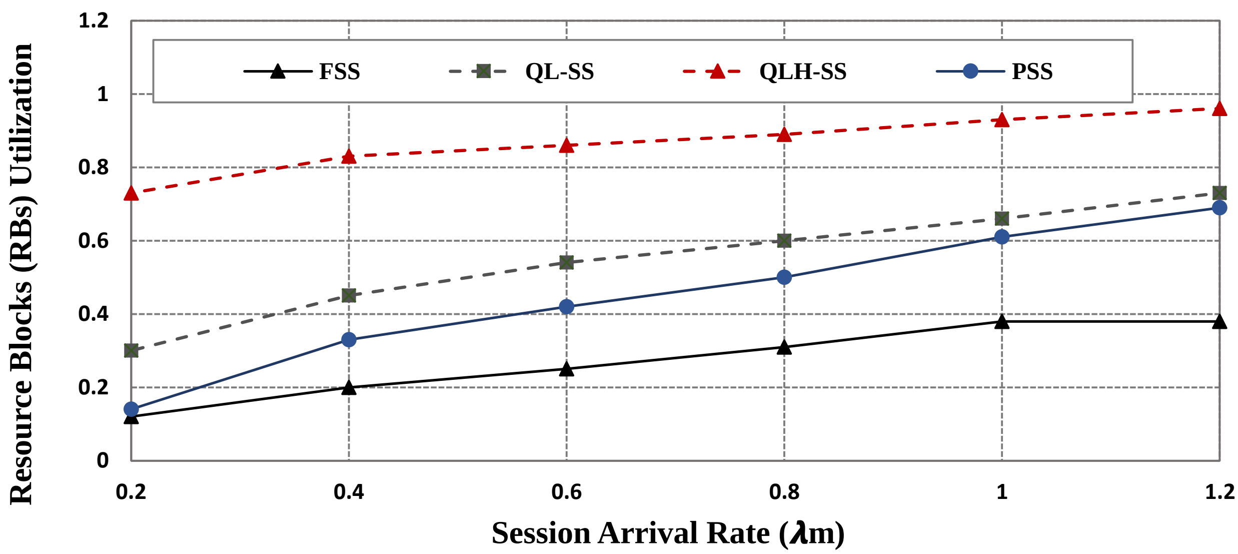


**Figure 3.6:** Downlink RB utilization as a function of the eMBB session generation rate $\lambda_m$ (sessions/s).

in uplink (i.e. QLH-SS achieves a relative gain of 82 %) and 39 % in downlink (i.e. QLH-SS achieves a relative gain of 146 %) when eMBB arrival rate is 1.2 sessions/s.

Fig. 3.7 presents the cumulative distribution function (CDF) of overall resource pool utilization for uplink and sidelink transmissions at arrival rate of 1.2 sessions/s. The results show that our proposed QL-SS and QLH-SS approaches maintains high resource utilization and outperforms the FSS and PSS reference schemes.

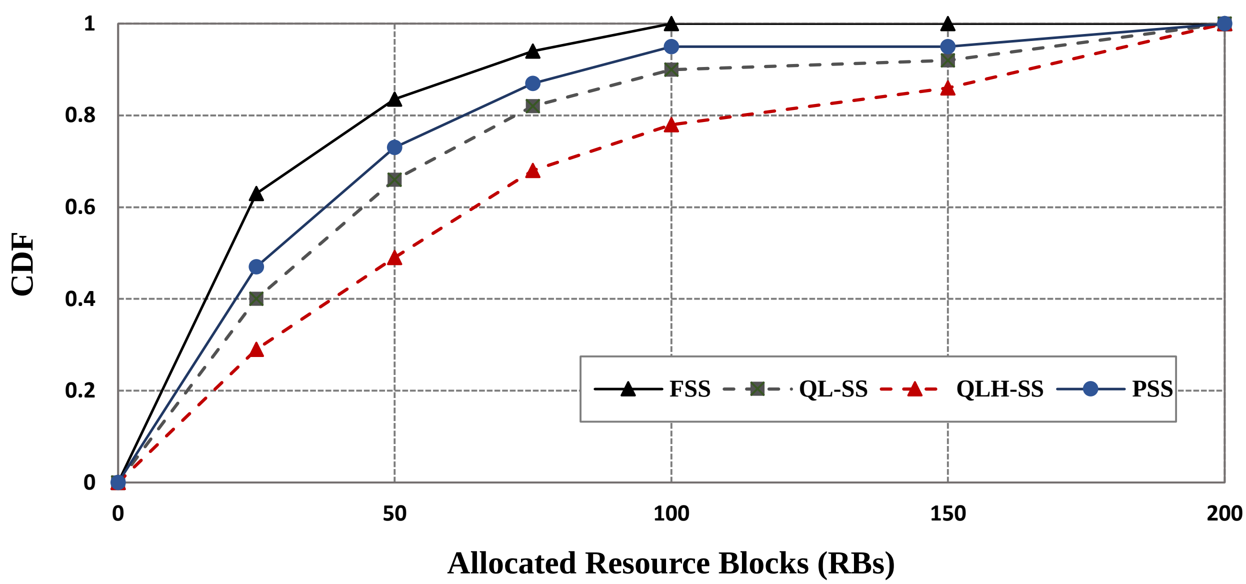


**Figure 3.7**: CDF of the number of utilized Resource Blocks by both slices in uplink and sidelink.

### 3.5.3.2 Performance in terms of Network Throughput

Figure 3.8 and Figure 3.9 depict the aggregate throughput delivered in Mbits/sec for both eMBB and V2X slices in the uplink (including both sidelink and uplink traffic) and downlink, respectively. The figures illustrate the behavior of the proposed solutions and the reference schemes. Here, we can observe that the proposed QL-SS and QLH-SS approaches outperform the references scheme. In the case of the QLH-SS, the maximum throughput for both services is about 136 Mb/s in uplink and 148 Mb/s in downlink, when the eMBB arrival rate is 1.2 sessions/s. In turn, the QL-SS achieved a maximum throughput of 118 Mb/s and 98 Mb/s in uplink and downlink respectively.

As shown by the results, the PSS approach achieved maximum throughput of 112 Mb/s in uplink (i.e. QLH-SS achieves a relative gain of 21 %) and 94 Mb/s in downlink (i.e. QLH-SS achieves a relative gain of 57%). Whereas in case of the PSS, the maximum throughput is only 74 Mb/s in uplink (i.e. QLH-SS achieves a relative gain of 83 %) and 60 Mb/s in downlink (i.e. QLH-SS achieves a relative gain of 146%). The reason for this behavior is that, as the number of eMBB sessions increases, requiring more radio resources, the QLH-SS ensures more RBs and achieves a higher utilization for radio resources than the reference schemes. Therefore, these RBs can be used to transmit data, while in the reference approaches there are no more available RBs for use in data transmissions.

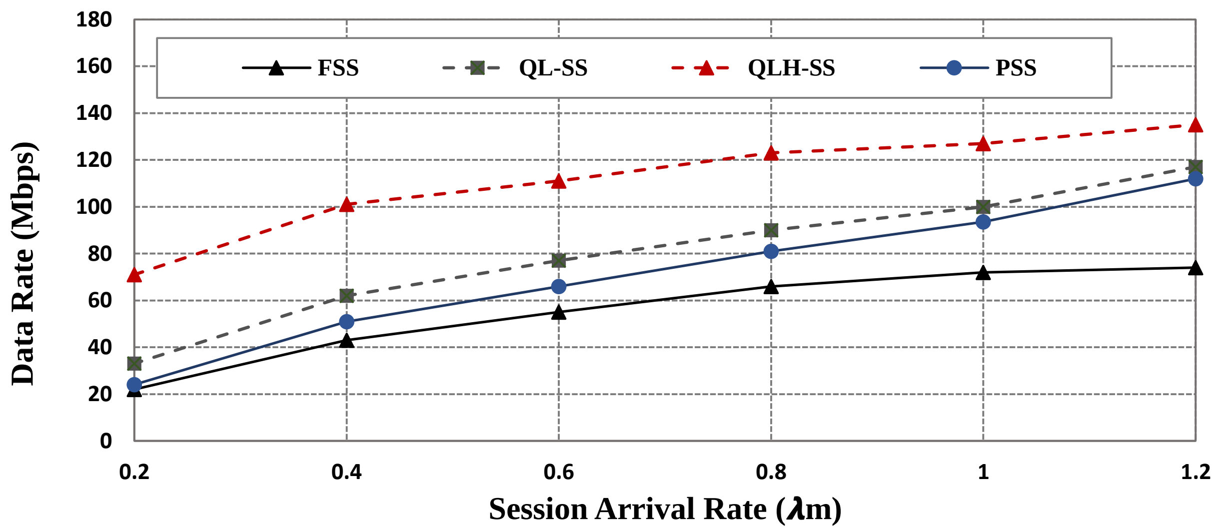


**Figure 3.8**: Aggregated throughput experienced by both slices in uplink (including uplink and sidelink traffic) as a function of the eMBB session generation rate $\lambda_m$ (sessions/s).

Figure 3.10 presents the CDF of the aggregate throughput achieved by the eMBB and V2X slices for both uplink and sidelink transmissions. As shown by the results, the proposed QLH-SS maintains a higher throughput than the PSS and FSS reference schemes. Specifically, the QLH-SS achieves up to 143 Mb/s with probability of 95 % while the PSS achieves 100 Mb/s and the FSS achieves only 80 Mb/s. This evaluation demonstrates the capability of the proposed schemes to utilize more RBs in order to achieve high data rate.

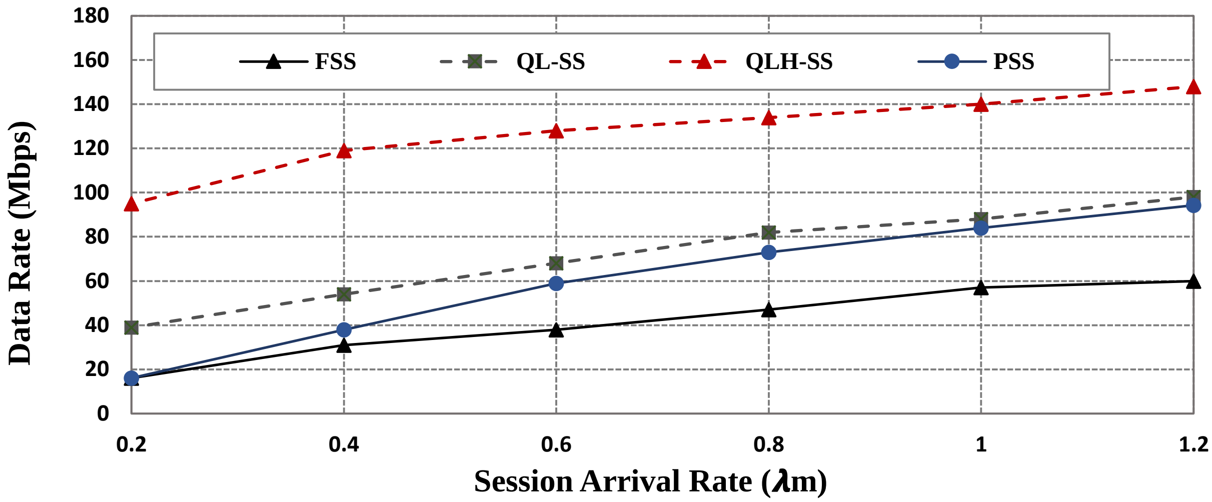


**Figure 3.9:** Aggregated throughput experienced by both slices in Downlink as a function of the eMBB session generation rate λm (sessions/s).

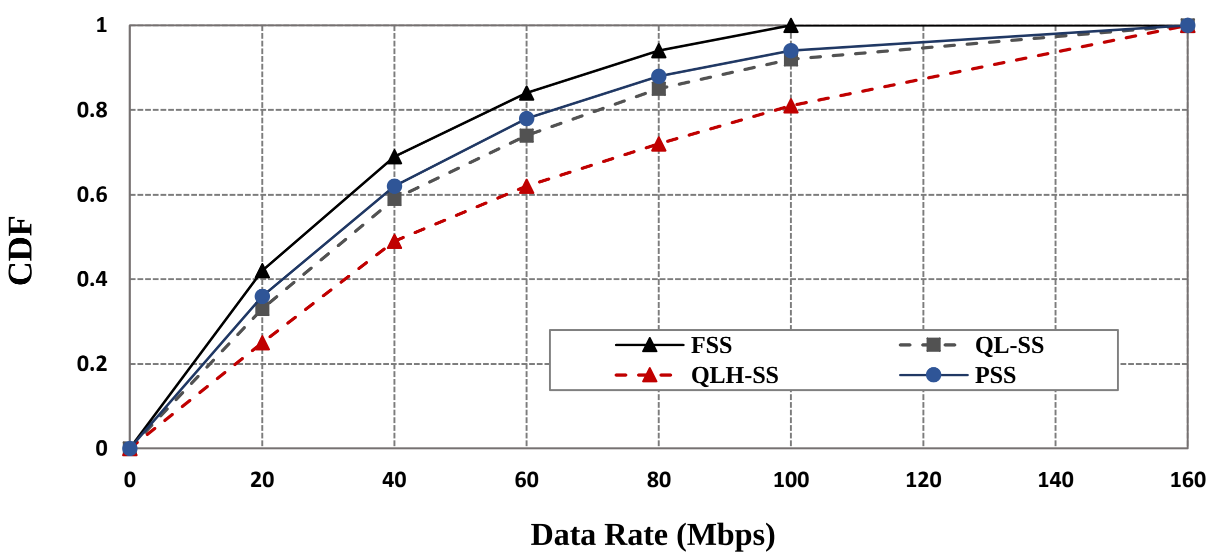


**Figure 3.10**: CDF of Aggregated throughput experienced by both slices in uplink and sidelink.

### 3.5.3.3 Performance in terms of Outage Probability

In Figure 3.11, we investigate the probability of having outage due to the lack of radio resources at a certain point of time. The outage probability of QL-SS , QLH-SS, PSS , and FSS schemes is a certain point of time. The outage probability of QL-SS , QLH-SS, PSS , and FSS schemes is plotted against the eMBB session arrival rate $\lambda_m$. As shown in the figure, increasing the traffic load leads to an increase in the outage probability. It can be also noted that our proposed QL-SS and QLH-SS schemes can substantially reduce the outage probability compared to the PSS and FSS reference schemes.

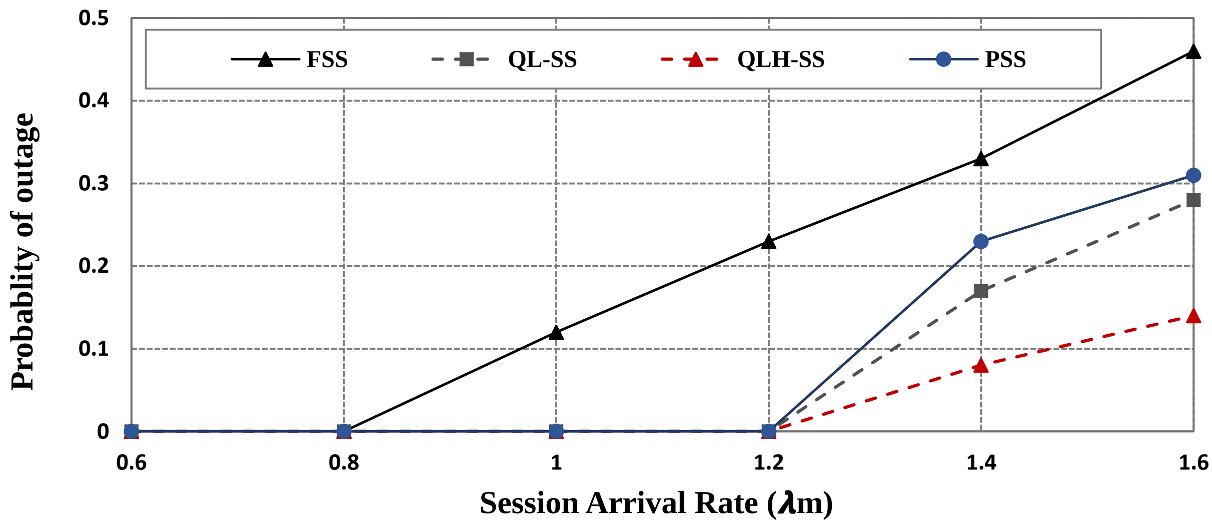


**Figure 3.11**: Outage probability as a function of the eMBB session generation rate λm (sessions/s).

### 3.5.3.4 Performance in terms of Latency

Figure 3.12 depicts the average latency for V2X service caused by channel access delay and the transmission delay. We can clearly observe that when vehicle arrival rate $\lambda_v$ is increased, more vehicles will use the network and request RBs to be used for the transmissions. This causes an increase in the waiting time and therefore increases the latency. We notice that the proposed QL-SS and QLH-SS reduces the latency compared to the PSS and FSS references schemes. For the QL-SS , when the packet arrival rate of V2X traffic is 10 packets/s, the average latency around 0.28s. Further improvements are obtained by the QLH-SS as the latency only about 0.13s. For the PSS approach, the latency is 0.31s In case of FSS, the latency is about 0.39s. This is due to the fact that the proposed solution guarantees higher availability of resources avoiding outage situations.

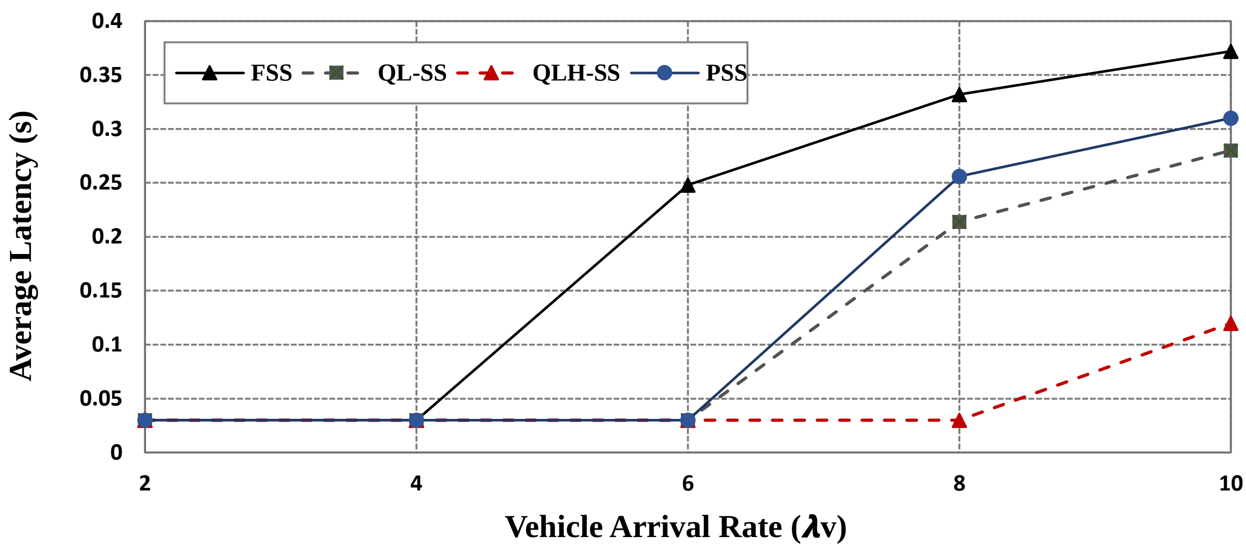


**Figure 3.12**: Average Latency as a function of the V2X UEs vehicle arrival rate $\lambda_v$ (vehicles/s).

### 3.5.4 Action Selection Probability

This section intends to analyse the behavior of the Q-learning scheme when selecting the different actions. For that purpose, Figure 3.13 presents the time for convergence, as a function of the number of actions. It is measured as the number of simulated time steps of 0.1s in the execution of the off-line RL until reaching convergence. From Figure 3.13, it is clearly observed that, as the number of actions increases, the number of time steps needed by the RL algorithm to converge also increases. This is because the system will try and explore more actions before converging to the selected one. Figure 3.14 and Figure 3.15 show two examples with the evolution of the probability of selecting each action $a_{k,x}$ in uplink and downlink, respectively. In the conducted simulations, the Q-learning approach with $\tau = 0.1$ is considered to control the tradeoff between exploration and exploitation in the learning mechanism.

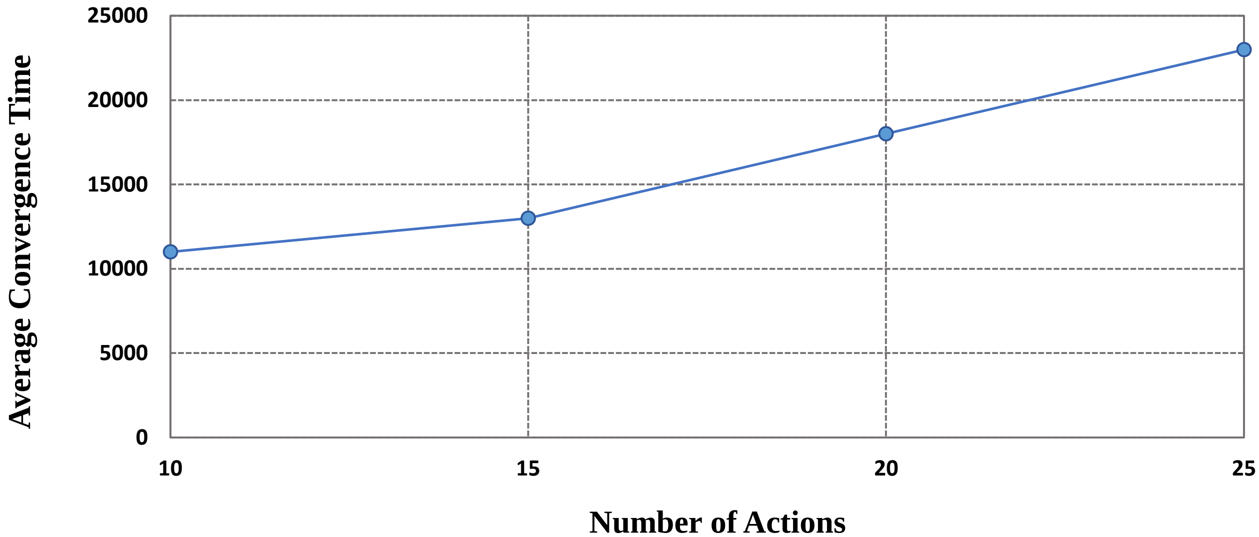


**Figure 3.13**: Convergence time as a function of Number of Actions.

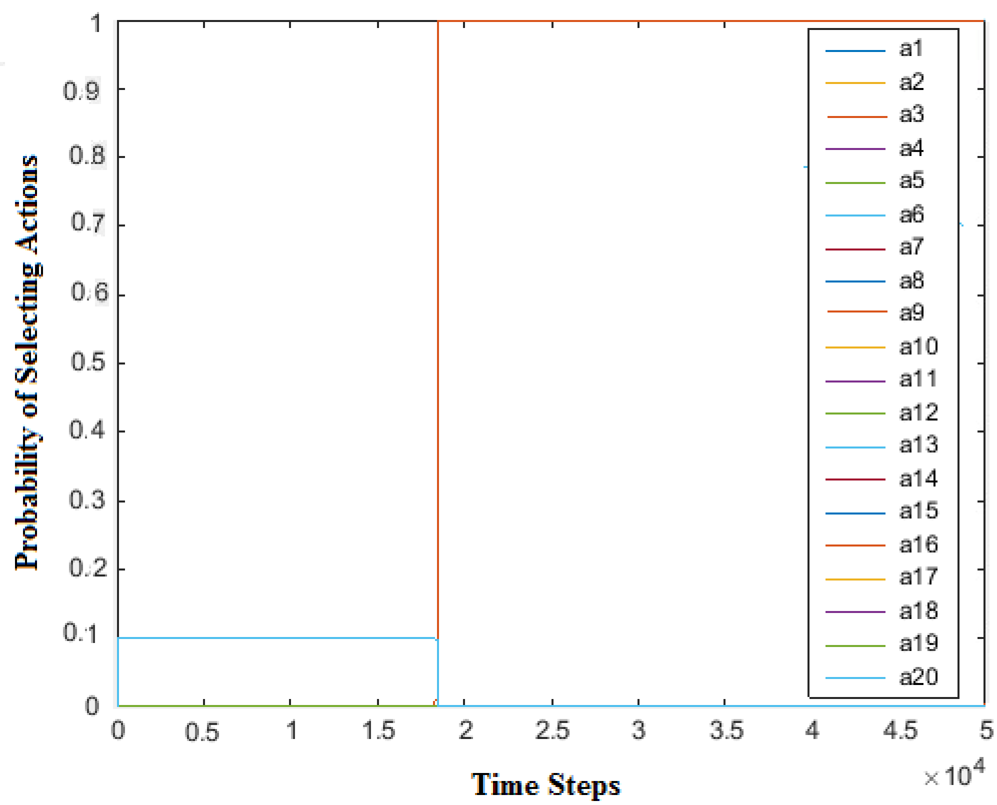


**Figure 3.14**: Selection probabilities $P(a_{k,UL})$ for the Uplink

As can be seen in Figure 3.14, at the beginning of the simulation, the proposed reinforcement learning starts to check different actions in order to choose the actions that yield most long-term reward. For the uplink, action $a_{3,UL}$, corresponding to $\alpha_{1,UL} = 0.15$ and $\alpha_{2,UL} = 0.85$, begins to increase and after some iterations and at approximately t = 18000 time steps, it is selected with probability close to 1 as the best action for determining the radio resources to each slice in uplink. Similarly, it can be seen in Fig. 3.15 that for the downlink, action $a_{1,DL}$, corresponding to $\alpha_{1,DL} = 0.05$ and $\alpha_{2,DL} = 0.95$, is selected as the best option. This occurs at approximately t = 6000 time steps.

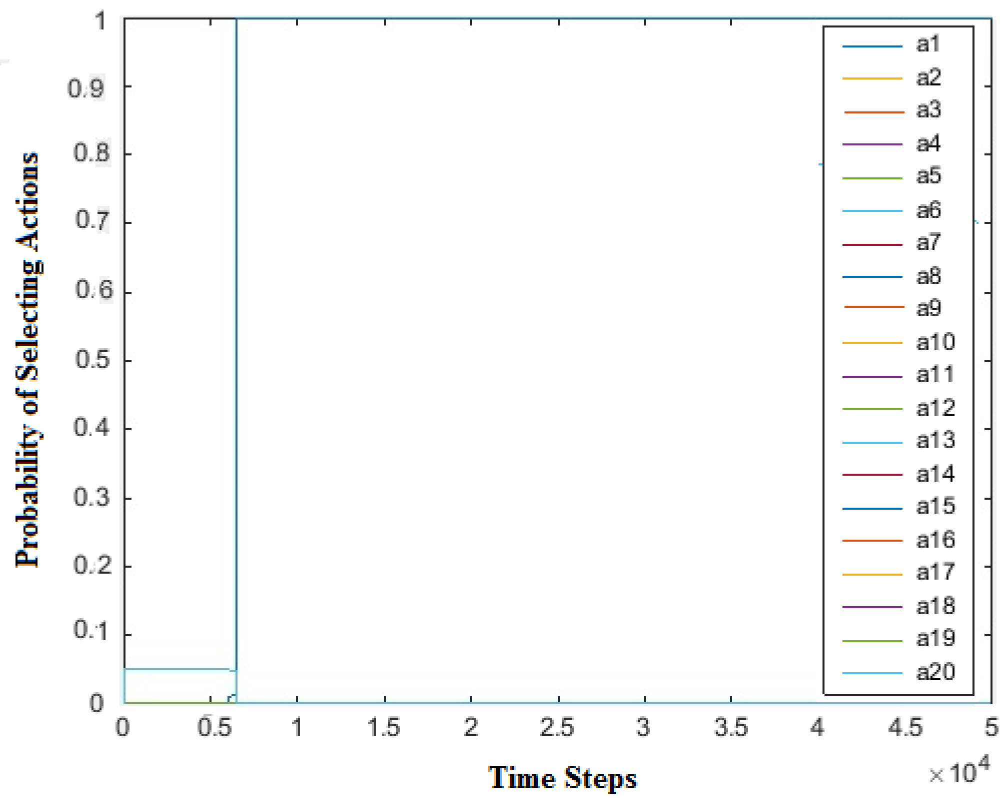


**Figure 3.15**: Selection probabilities $P(a_{k,DL})$ for the Downlink.

## 3.6 Concluding Remarks

In this chapter, we have investigated the problem of how to split the radio resources between multiple RAN slices in a scenario with V2X and eMBB services involving uplink, downlink and sidelink (for direct V2V) communications. We have proposed a new RAN slicing strategy based on an off-line reinforcement learning followed by a low-complexity heuristic approach to determine the split of resources assigned to the eMBB and V2X slices.

Extensive simulations were conducted to validate and analyse the performance of our proposed solution, comparing it against two reference schemes and against the case with only the RL approach. Simulation results show the capability of the proposed algorithms to allocate the resources efficiently and improve the network performance in terms of resource utilization, throughput, latency and outage probability. Specifically, from the presented results, we notice that our proposed scheme including both the off-line Q-learning and the low-complexity heuristic approach outperforms the reference based on a distribution proportional to the traffic and the reference with fixed slicing ratio in terms of the achieved Resource Block Utilization with gains of up to 32 % and 82 %, respectively, in uplink. Results also demonstrated that bit rate improvements of up to 21% and 83% can be obtained in uplink with respect to the case with off-line Q-learning followed by a low-complexity approach compared to the reference with proportional approach and the reference with fixed slicing strategy, respectively. The proposed algorithms can reduce latency caused by channel access delay and the transmission delay for V2X communications (i.e. latency reductions of around 0.18s and 0.26s have been obtained with the proposed approach with respect to the reference based on a distribution proportional to the traffic and the reference with fixed slicing ratio, respectively). Moreover, the proposed algorithms can also reduce the outage probability that arises when increasing the traffic and thus the demands for RBs (i.e. proposed approach reduces the outage probability by around 0.16 and 0.32 with respect to the reference based on a distribution proportional to the traffic and the reference with fixed slicing ratio, respectively).

# Chapter 4

# Mode Selection for V2V communications in Cellular Networks

## 4.1 Introduction

SL communication underlying cellular networks is a new paradigm that has been proposed to enhance the performance of cellular networks and support V2X communications. Integration of SL in cellular network allows a UE to operate either in cellular mode, based on uplink/downlink communications, or sidelink mode, which enables vehicles to communicate directly with other vehicles in close proximity over the PC5 interface. By properly selecting the operating mode according to the requirements of V2X services, significant operational benefits and efficient utilization of the spectrum can be achieved. However, selecting the appropriate operating mode according to the requirements of V2V services becomes a challenging issue. In this respect, two different solutions for mode selection are proposed in this chapter for V2V communication over a cellular network. In the first solution, a novel mode selection strategy to decide when it is appropriate to use one or the other mode for the involved vehicles, taking into account the quality of the links between V2V users in sidelink mode and between the base station and the vehicles in cellular mode, the available resources, and the network traffic load situation. Moreover, in order to achieve further improvements in the network performance (i.e., in terms of resource utilization, the latency of V2X services, achievable data rate, and outage probability), another mode selection and resource reuse strategy is proposed to select the appropriate mode of operation and decide the amount of resources to be allocated to V2V links in each mode. Different from the first mode selection solution, the second mode selection strategy allows reusing RBs between different SL users, provided that the interference constraints are met.

The remainder of this chapter is organized as follows. Section 4.2 introduces the state of the art and summarizes the contributions of the chapter. Section 4.3 presents the system model along with the assumptions. Section 4.4 discusses the first mode selection solution that takes into account the quality of the links between V2V users in sidelink mode and between the base station and the vehicles in cellular mode, and the available resources. In addition, the second mode selection strategy that exploits resource reuse is introduced in Section 4.4, along with the performance evaluation for the proposed strategies. Finally, conclusions are given in Section 4.5.

## 4.2 State of the Art and contributions

The widely deployed cellular network, assisted with device to-device (D2D) communications, can provide a promising solution to support V2V communications. There have been different considerable works that have been presented to explore mode selection and resource optimization in direct Device-to-Device (D2D) communications taking into account the interference situation [66-72] when sharing the radio resources and the quality of the links for both cellular and D2D links [73- 78]. In [66], joint mode selection, channel assignment and power control in D2D communications are investigated. They proposed low-complexity algorithms aiming at maximizing the overall system throughput while guaranteeing the signal-to noise-and-interference ratio of both D2D and cellular links according to different network loads. A holistic approach for D2D mode selection and interference alignment technique for interference management is proposed in [67]. Reference [68] proposed a joint D2D mode selection and resource allocation scheme in order to maximize the system sum rate while meeting the successive interference cancellation (SIC) decoding constraint. The authors of [69] investigated the joint mode selection and channel assignment in a cellular network with underlying D2D communications, where multiple D2D links may share the same channel. Meanwhile, the QoS requirements for both D2D links and cellular users are guaranteed, in terms of Signal-to-Interference-Plus-Noise Ratio (SINR). The authors in [70] proposed a scheduling algorithm to enable the eNB to perform joint scheduling on mode selection, radio resource allocation, and power coordination in D2D communication underlying cellular networks. A novel joint mode selection and channel resource allocation algorithms via the vertex coloring approach are proposed in [71]. Reference [72] proposed a network-assisted mode selection mechanism in a multi-cell scenario when taking into account the QoS requirements for cellular users and resource sharing as well as power optimization for D2D users. However, most of the existing solutions in the literature focused on maximizing the sum rate of all users [66], [68], [69], [71]. Whereas V2X services are event triggered or periodic, the size of Basic Safety Message (BSM) / CAM is usually small, so developing strategies to deliver the messages quickly, balancing the network load and offloading the network to avoid deterioration in the quality of services (QoS), and achieve significant system improvements in terms of resource consumption are important issues for V2X applications and need to be investigated to ensure that all the vehicles within the network will receive a high QoS evenly.

A new mode selection scheme for device-to-device (D2D)-enabled cellular communications with mobility based on an average threshold D2D distance between two given users is proposed and evaluated in [73]. In [74], a mode selection approach based on evolutionary game is proposed in UAV-aided vehicular network, and an evolutionarily stable strategy is obtained. Similarly, [75] proposed a mode selection strategy to determine the mode of operation and allow D2D pairs to flexibly reuse the radio resources of cellular users to improve the quality of D2D links. The authors in [76] studied the mode selection problems in a multi-mode and multi-pair D2D network, where the eNB can assign one of the three D2D communication modes including local route mode, direct D2D mode, and relay D2D mode. An online learning technique which leverages combinatorial multi-armed bandits (CMAB) is proposed in [77] to tackle the combinatorial nature of the mode selection and resource allocation (MS&RA). In turn, [78] proposed a dynamic mode selection and subchannel allocation for an orthogonal frequency-division multiple access (OFDMA) cellular network with D2D communications to minimize the average end-to-end delay performance under the dropping probability constraint. However the mode selection schemes in [73]–[78] are conducted for conventional D2D systems with users at slow mobility and not for V2X communications.

Based on all the above, the main contributions of this chapter are the proposal and analysis of different solutions for mode selection for V2V communication over the cellular networks. In the first solution, we propose a novel mode selection strategy that takes into account the quality of the links between V2V users in sidelink mode and between the base station and the vehicles in cellular mode, the available resources, and the network traffic load situation. The proposed mode selection strategy is introduced in order to precisely achieve the following objectives: (i) minimizing the consumption of radio resources (ii) reducing the network traffic congestion, and (iii) ensuring a high quality of signal for all users in the network without causing significant harm to the cellular network.

In the second solution, a novel mode selection and resource reuse strategy to select the appropriate mode of operation and decide the amount of resource blocks (RBs) for V2V links in cellular and sidelink modes, where multiple V2V links may share the same radio resources. Such a strategy would be useful in order to face different challenges in terms of capacity, latency, reliability, and scalability caused by the increase in network size and demands for radio resources. Moreover, the impact of mode selection deployment on the performance of the underlying cellular network is extensively studied in this chapter via simulations.

## 4.3 System Model and Assumptions

The considered scenario assumes a cellular Next Generation Radio Access Network (NG-RAN) with different gNodeBs (gNBs) deployed along a highway. A roadside unit (RSU) supporting V2X communications is attached to each gNB. A flow of several independent vehicles move along a straight highway, as illustrated in Figure 4.1. The highway segment is divided into sub-segments (clusters) by sectioning the road into smaller zones according to the length of the road. It is assumed that each vehicle includes a User Equipment (UE) that enables communication with the UEs in the rest of vehicles in the same cluster. Clusters are numbered as $j$=1,...,$C$, and the vehicles in the $j$-th cluster are numbered as $i$=1,...,$V(j)$. We assume that V2V communication between vehicles can be performed either in cellular or in sidelink mode.

In cellular mode, each UE communicates with each other through the Uu interface in a two-hops transmission (i.e. uplink and downlink) via the gNB while in sidelink mode, direct V2V communications can be established over the PC5 interface. We introduce an indicator $\alpha_j$ to reflect the operation mode of the $j$-th cluster, so that $\alpha_j$ =1 if vehicles within the cluster operate in sidelink

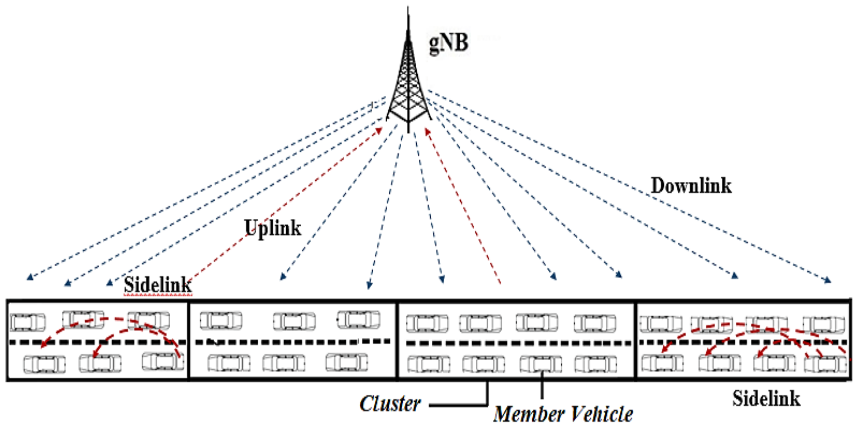


**Figure 4.1**: System model of the cellular network with sidelink V2V.

mode and $\alpha_j = 0$ if they operate in cellular mode. We assume that, when sidelink transmissions are utilized, every member vehicle can multicast the V2V packets directly to multiple member vehicles of the same cluster $1 \leq i \leq V(j)$ using one-to-many technology.

Concerning the V2V clustering, key benefits of this concept when it is integrated into the cellular system are:

- By dividing the network into clusters, the consumption of the resources will be reduced especially when the density of the network is high. Vehicles have to exchange a high number of messages and therefore they create additional traffic data and consume a high quantity of resources especially when each member vehicle needs to communicate with the gNB directly [79].

- Guaranteeing the required latency and reliability for V2X services is challenging. The requirement for V2X communication on high reliability is congruent with the proximity gain provided by the sidelink and partitioning devices into groups (clusters) is necessary to achieve the proximity. Sidelink clustering concept is considered to reduce the latency, especially in the cases when the gNB is far away from the vehicles. If the transmitting vehicles transmit the safety messages to other members, it is feasible to mimic the scheduled transmissions in a cell and improve the system efficiency [80].
- Using clustering based location (i.e., by grouping the vehicles into clusters based on location), the same radio resources of V2V links within a cluster in SL can be reused by

other V2V links within other clusters in SL without interfering with each other, and this will significantly improve system capacity, spectral efficiency. Flexible and inflexible cells clustering-based interference mitigation schemes are summarized in [81,82].

The vehicles in the highway are assumed to enter the cell coverage following a Poisson process with arrival rate $\lambda_v$. The association between clusters and vehicles is managed and maintained by the RSU based on different metrics (e.g., position, direction, speed and link quality) through a periodic exchange of status information. The number of available RBs in UL and DL are denoted as $N_{UL}$ , $N_{DL}$ , respectively. Regarding the radio resources for V2V in uplink and sidelink, we assume that, like in LTE, the sidelink RBs are a subset of the uplink RBs and all clusters in SL get the resource blocks from UL resource blocks.

## 4.4 Mode Selection Strategies

In this section, we present two different novel mode selection strategies to decide when it is appropriate to use one or the other mode for V2V-enabled cellular networks. In the first strategy, referred to as mode selection based signal strategy (MSBS) we consider a mode selection algorithm takes into account the quality of the links between V2V users in sidelink mode and between the base station and the vehicles in cellular mode, and the available resources. In addition, in order to achieve further improvements in the use of resources, latency, and congestion, another novel mode selection strategy that exploits resource reuse is introduced, where multiple V2V links may share the same radio resources. This second strategy is referred to as mode selection and RBs reuse strategy (MS-RBRS).

### 4.4.1 Mode Selection based Signal Strategy (MSBS)

#### 4.4.1.1 Description of the MSBS Strategy

In this section, we introduce the key steps for the proposed mode selection based signal strategy (MSBS) in detail. The proposed strategy operates periodically in time windows of duration *T* and is detailed in the pseudo-code of Algorithm 4.1. The proposed strategy assumes that cluster j is working in mode $\alpha_j$ at a given time window, and the algorithm determines the mode for the next time window. To make this decision, the proposed mode selection strategy considers the following conditions for cluster *j*:

1- It must ensure high signal quality by connecting V2V links to the mode with better signal-to-interference plus-noise ratios (SINRs). Let $\gamma_{j,i}^{UL}$ (t) and $\gamma_{j,i}^{DL}$ (t) denote the measured SINR for the *i*-th vehicle in the uplink and downlink of cellular mode, respectively, and $\gamma_{j,i,i'}^{SL}$ the SINR in the sidelink between the *i*-th and the *i*'-th vehicles of the *j*-th cluster. Each vehicle transmits the information about the received SINR from the gNB and other transmitting vehicles respectively to the gNB. Then, the average SINR for each cluster of vehicles is computed for each mode of operation and compared to each other (lines 4, 5).

The average SINR for all the UEs in uplink and downlink transmissions within each cluster in a time window T can be statistically estimated as follows:

$$\overline{\gamma}_j^{x} = \frac{1}{T.V(j)} \sum_{t=1}^{T} \sum_{i=1}^{V(j)} \gamma_{j,i}^{x}(t) \tag{4.1}$$

where $x \in \{UL,DL\}$. As for the sidelink mode, since multicast technology is utilized to transmit the packets from vehicle *i* to other vehicles within the cluster for $1 \le i' \le V(j)$, $i' \neq i$, the average of SINR for all the UEs within cluster j in a time window T is calculated as:

$$\overline{\gamma}_j^{SL} = \frac{1}{T.V(j).(V(j)-1)} \sum_{t=1}^{T} \sum_{i=1}^{V(j)} \sum_{\substack{i'=1 \\ i'\neq i}}^{V(j)} \gamma_{j,i,i'}^{SL}(t) \tag{4.2}$$

2- The V2V links to be established in sidelink or cellular must ensure that the amount of required RBs by cluster j is less than the number of RBs available for the mode in which the vehicles are to be switched and operated. i.e., for every j ∈ C, the following condition is checked (lines 6, 10):

**Algorithm 4.1 :** MSBS Algorithm

1. **Inputs:**
   $N_{UL}$, $N_{DL}$: Number of RBs in UL and DL.
   C: number of clusters.
   V(j): vehicles in cluster j,
2. **Initialization**: $\rho^{x}(t)=0$
3. **For** each cluster $j \in C$
4. Compute the average value of $\bar{\gamma}_j^{UL}, \bar{\gamma}_j^{DL}$ , $\bar{\gamma}_j^{SL}$ among all vehicles within cluster j using eqs. (4.1) and (4.2);
5. If $Min\left(\bar{\gamma}_j^{UL}, \bar{\gamma}_j^{DL}\right) \geq \bar{\gamma}_j^{SL}$
6. If $\Gamma_j^x \leq N_x - \rho^x$ {check that the required RBs by cluster j are available in each link $x \in [UL, DL]$}.
7. $\alpha_j=0$ cluster j operates in cellular mode
8. $\rho^{UL} = \rho^{UL} + \Gamma_j^{UL}$
9. $\rho^{DL} = \rho^{DL} + \Gamma_j^{DL}$
10. Else If $\Gamma_j^{SL} \leq N_{RB}^{UL} - \rho^{UL}$ { check that the required RBs by cluster j are available in UL
11. $\alpha_j=1$ { cluster j operates in sidelink mode}
12. $\rho^{SL} = \rho^{SL} + \Gamma_j^{SL}$
13. End
14. Else
15. $\alpha_j=1$ {cluster j operates in sidelink mode}
16. $\rho^{SL} = \rho^{SL} + \Gamma_j^{SL}$
17. End
18. **End**
19. **Output:** $\alpha_j$, $j=1,..,C$

$$\Gamma_j^x < N_x - \rho^x(t) \tag{4.3}$$

where $N_x$ is total the number of RBs in the link $x \in$ {UL,DL}. $\rho^x$ is the total number of RBs that have been allocated to link $x \in$ {UL,DL,SL}and determined by algorithm 4.1 (lines 8,9,12,16). The average is measured over time window *T*. $\Gamma_j^x$ is the average number of required RBs from V2X users in cluster j in UL, DL and SL over the time window *T*, given by :

$$\Gamma_j^x = \frac{\sum_{t=0}^{T}\sum_{i=1}^{V(j)} m(j,i,t) \cdot S_m}{T \cdot SP_{eff,x} \cdot B \cdot T_s} \tag{4.4}$$

where *x* denotes the type of link, i.e. $x \in \{UL,DL,SL\}$, m(*j*,*i*,*t*) is the number of transmitted packets by the vehicles of the j-th cluster at each time *t* within the time window T, and $SP_{eff,x}$ is the spectral efficiency associated to the modulation and coding scheme to be used in the x link, $T_s$ is the symbol duration and B the bandwidth per RB.

Based on the above conditions, the mode selection criterion is as follows: If the average values of SINR for both uplink and downlink in the j-th cluster are higher than the average value of sidelink SINR, (i.e. $\min(\bar{\gamma}_j^{UL}, \bar{\gamma}_j^{DL}) > \bar{\gamma}_j^{SL}$) , and the number of available RBs in each link $x \in [UL, DL]$ is higher than the RBs required by V2V links within the cluster, then it is switched to the cellular mode, otherwise the cluster will stay in sidelink mode as lines (5-17) show.

#### 4.4.1.2 Performance Evaluation for MSBS

In order to demonstrate the feasibility of the proposed mode selection algorithm, a single cell is simulated in MATLAB. All relevant system and simulation parameters are summarized in Table 4.1.

The main characteristic of the proposed MSBS strategy is to reduce the channel congestion due to an increase in demands for the RBs and achieve an efficient usage of the available RBs at the cell. Then, in order to get a vision and check the behavior of the algorithm, different performance metrics such as offered load, required RBs, and traffic congestion are evaluated. The MSBS strategy is compared against a reference scheme in which the vehicles always operate in sidelink mode.

**TABLE 4.1:** System Parameters used in the Simulation of MSBS.

<table>
<tr><th>Parameter</th><th colspan="2">Values</th></tr>
<tr><td rowspan="3">Number of RBs per cell</td><td rowspan="2">100 RBs<br>For uplink pool</td><td>50% for sidelink transmission (sidelink mode)</td></tr>
<tr><td>50% for uplink transmission (cellular mode)</td></tr>
<tr><td colspan="2">100 RBs for downlink broadcast (cellular mode)</td></tr>
<tr><td>Number of lanes</td><td colspan="2">3 in one direction (one is considered in the freeway)</td></tr>
<tr><td>Length of the street</td><td colspan="2">Freeway length = 1 Km</td></tr>
<tr><td>Lane width</td><td colspan="2">4 m</td></tr>
<tr><td>Size of cluster</td><td colspan="2">250m</td></tr>
<tr><td>Number of clusters</td><td colspan="2">4</td></tr>
<tr><td>Vehicular UE height</td><td colspan="2">1.5m</td></tr>
<tr><td>Packet size (Sm)</td><td colspan="2">{300,800,1200} bytes</td></tr>
<tr><td>Packet generation rate</td><td colspan="2">1 [packets/s]</td></tr>
<tr><td>Path loss model</td><td colspan="2">The path loss and the LOS probability for cellular mode are modeled as in [63]. In sidelink mode, all V2V links are modeled based on freeway case (WINNER+B1) with hexagonal layout [ITU-R] [64].</td></tr>
<tr><td>Spectral efficiency model to map SINR.</td><td colspan="2">Model in section A.1 of [65]. The maximum spectral efficiency is 4.4 b/s/Hz.</td></tr>
<tr><td>Base station antenna gain</td><td colspan="2">5 dB</td></tr>
<tr><td>Shadowing standard deviation</td><td colspan="2">3 dB in LOS and 4 dB in NLOS.</td></tr>
<tr><td>Transmitted power per RB</td><td colspan="2">23 dBm</td></tr>
<tr><td>Frequency</td><td colspan="2">5.9 GHz</td></tr>
<tr><td>Time window T</td><td colspan="2">3s</td></tr>
<tr><td>vehicle speed</td><td colspan="2">80 Km/h</td></tr>
<tr><td>vehicle arrival rate ($\boldsymbol{\lambda}$v).</td><td colspan="2">Poisson model. Different average values considered in the simulations to simulate different loads.</td></tr>
</table>

Figure 4.2 depicts the data traffic load delivered in Kbits/sec for the sidelink communications with the MSBS strategy and with the reference scheme. The figure illustrates with three lines the

behavior for different packet sizes $S_m$ (300,800, and 1200) bytes. Here, we can clearly observe that when vehicle arrival rate $\boldsymbol{\lambda}_v$ is increased, more vehicles will use the network and request RBs to be used for the transmissions. For all the assumed arrival rates under consideration, the figure shows that there is a marked decrease in data traffic load for sidelink direct V2V system with respect to the reference scheme. This occurs because, with the MSBS strategy, some users switch

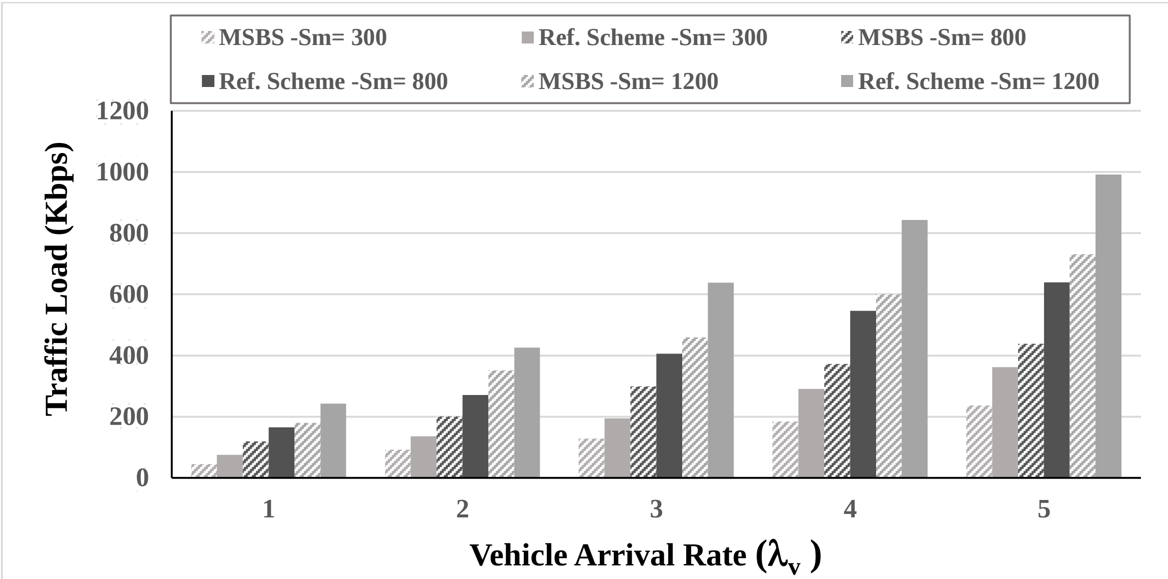


**Figure 4.2:** Network Traffic Load for sidelink V2V communications as a function of the vehicle arrival rate $\lambda_v$ (vehicles/s).

to operate and send the safety information through the cellular mode instead of through the sidelink mode. Figure.4.3 presents the RB utilisation for the sidelink (i.e. the number of used RBs normalized to the number of total available RBs for the SL). It shows that the MSBS strategy is able to reduce resource consumption. In the MSBS strategy, the required RBs for SL transmission are reduced when the arrival rate and thus the number of vehicles transmitting packets is reduced and when the packet sizes are small. By switching the vehicles who receive a high SINR from the gNB to operate in cellular mode, we keep some proportion of RBs to be utilized by the vehicles that do not receive a good quality of service from the gNB and are able to operate in SL mode only. From Figure 4.3, it is observed that our MSBS strategy saves up to 20 % of the SL resources with respect to the reference scheme.  This turns into an increase in the RB utilisation for cellular mode in both UL and DL, as shown in Figure 4.4 and Figure 4.5, respectively. Hence, the MSBS strategy ensures that all the vehicles within the network will receive a high QoS evenly, offloading the network and avoiding the outage of service due to the lack of resources. From Figure 4.4, it is observed that the maximum RB utilization in the uplink is about 48% when the vehicles arrival rate is 5 vehicles/s and the packet size is 1200 bytes, while for downlink transmission (see Figure 4.5), the maximum RB utilization is only about 24% when the vehicles arrival rate is 5 vehicles/s and the packet size is 1200 bytes. It is worth mentioning that Figure 4.4 and Figure 4.5 only depict the case of the MSBS strategy because with the reference scheme there are no vehicles transmitting in cellular mode.

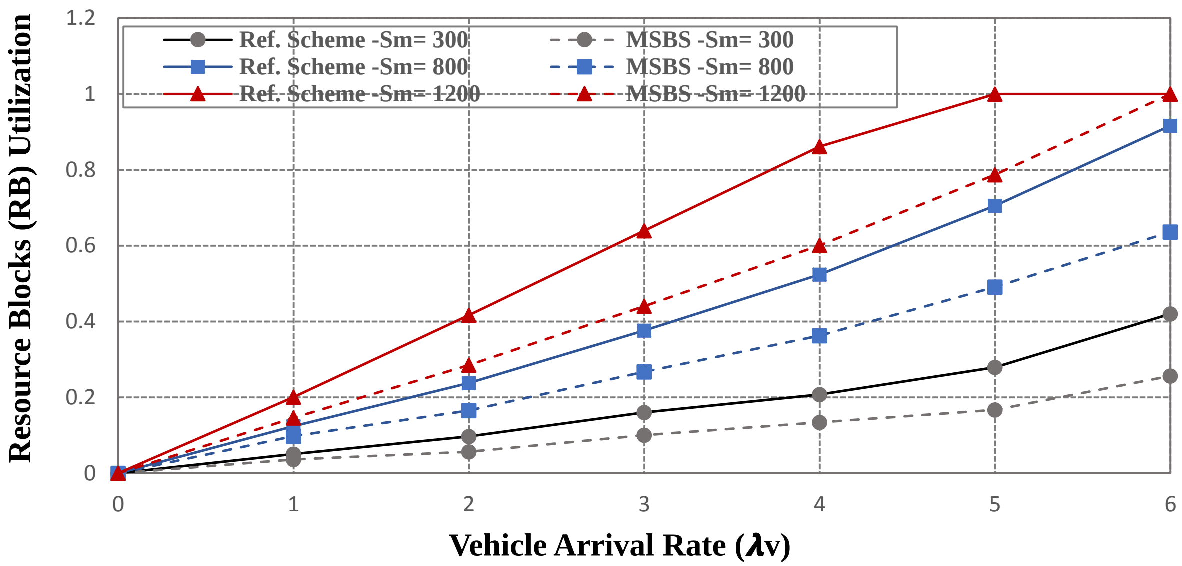


**Figure 4.3:** Utilization of Resource Blocks in sidelink mode as a function of the vehicle arrival rate $\lambda_v$ (vehicles/s).

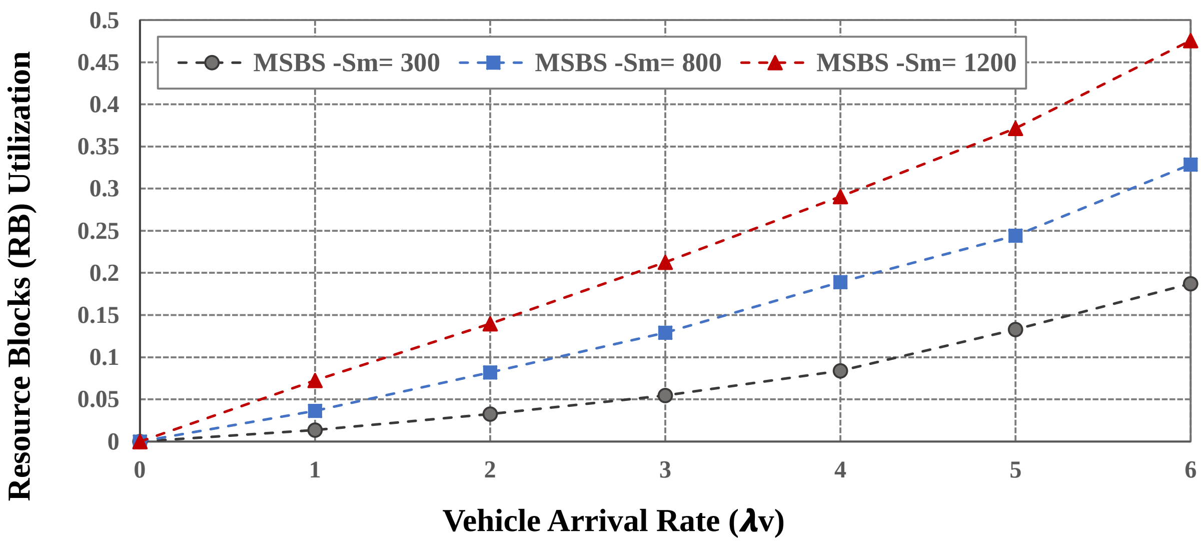


**Figure 4.4:** Resource Blocks Utilization for uplink transmissions as a function of the vehicle arrival rate $\lambda_v$ (vehicles/s)

Another relevant aspect is that the MSBS strategy can reduce the network traffic congestion result due to the increase in demands for the RBs. This is observed in Figure 4.6, which depicts

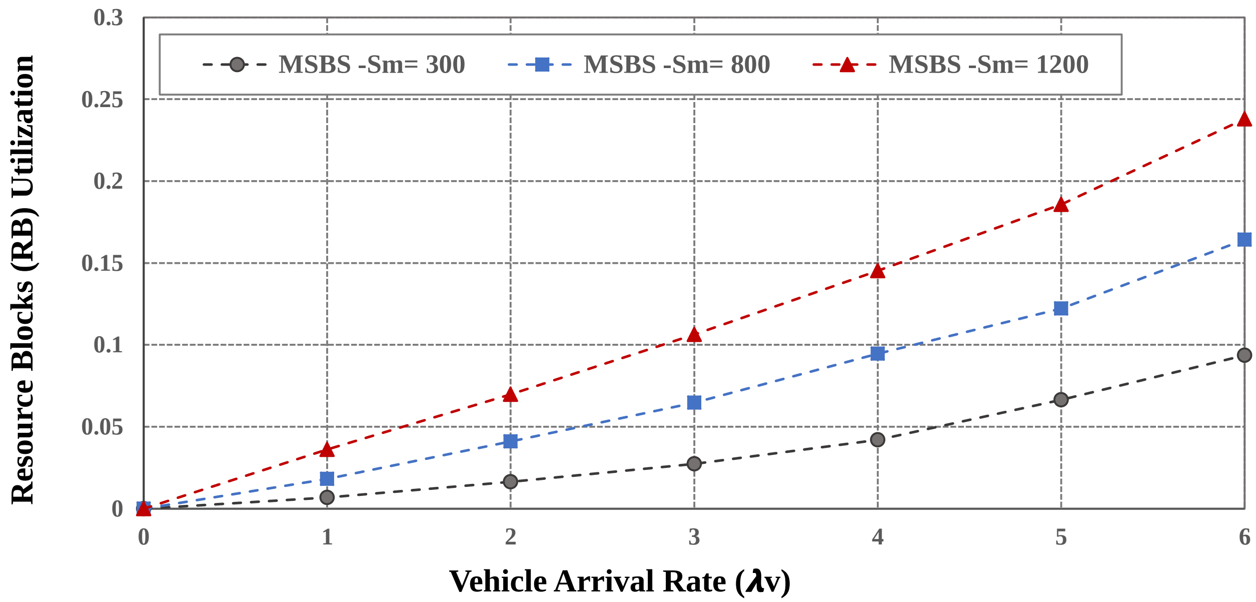


**Figure 4.5:** Resource Blocks Utilization for downlink transmissions as a function of the vehicle arrival rate $\lambda_v$ (vehicles/s)

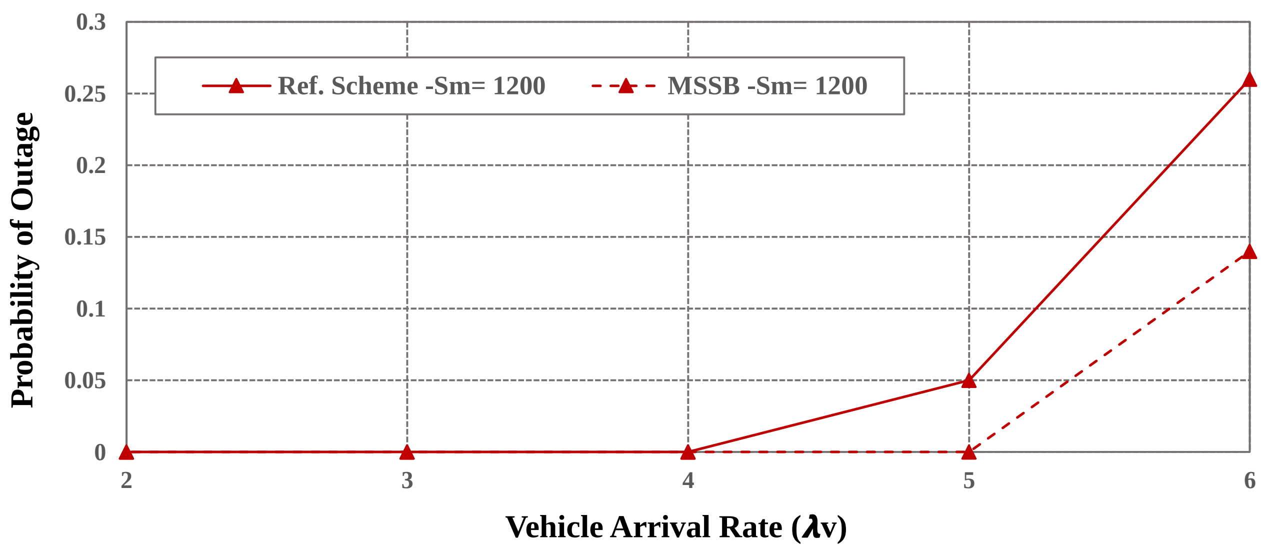


**Figure 4.6:** Probability of Outage in Sidelink as a function of the vehicle arrival rate λv (vehicles/s)

the probability of outage (i.e. the probability that there are not sufficient RBs to serve all the transmission requests). Results are presented for packet size 1200 bytes. It is observed that the MSBS strategy can substantially reduce the outage probability in relation to the reference scheme.

### 4.4.2 Mode selection based RBs reuse strategy (MS-RBRS)

#### 4.4.2.1 Description of the MS-RBRS Algorithm

This section considers a mode selection and RBs reuse strategy (MS-RBRS), in which the system keeps track of the performance of both modes, i.e. SL and CL, for each cluster in terms of the received signal and the availability of resources. The MS-RBRS strategy operates periodically in time windows of duration *T* and is detailed in the pseudo-code of Algorithm 4.2. Like in the MSBS approach, we assume an indicator $\alpha_j$ to reflect the operation mode of the *j*-th cluster, so that $\alpha_j$ =1 if vehicles within the cluster operate in sidelink mode and $\alpha_j$ =0 if they operate in cellular mode. The MS-RBRS strategy assumes that cluster j is working in mode $\alpha_j$ at a given time window, and the algorithm determines the mode for the next time window.

In the MS-RBRS strategy, three different aspects are considered:

First, it must ensure that the amount of required RBs that vehicles will need to transmit their packets is less than the number of available RBs in the network that are allocated for the mode in which the vehicles are to be switched and operated.

Second, ensure reliable connection by switching all the vehicles within the cluster to be served by the mode that has a higher link quality. In this algorithm, each vehicle transmits the information about the received SINR from the gNB and other transmitting vehicles respectively to the gNB. Then, the average SINR for each cluster of vehicles is computed for each mode of operation and compared to each other (lines 4, 5). The average SINR for all the UEs in sidelink, uplink, and downlink transmissions within each cluster in a time window T can be statistically estimated based on (4.1) and (4.2).

Third, in order for the V2V links within cluster j in SL mode to reuse other radio resources used by other V2V links in cluster $k$=1,...,$C$, $k \neq j$ operating in SL mode (cluster k with $\alpha_k$ =1), it must be ensured that the SINR between cluster j and cluster k is higher than a threshold , i.e $\gamma_{j,k} > \gamma_{TH}$. Since the positions of the transmitters and receivers will change dynamically, in order to compute the SINR between cluster j and cluster *k* on a long-term basis that is valid for any condition, let us assume the worst case situation between any i transmitter of cluster *j* and any receiver of cluster *k* (i.e., the vehicles positions that lead to lowest SINR). The worst case positions are denoted as WP1 and WP2 and shown in Figure. 4.7. Based on this, the SINR between cluster j and cluster k, $\gamma_{j,k}$ is calculated as follows :

$$\gamma_{j,k} = \frac{P_T G_T G_R}{L_{j,k} . (P_n + I_{j,k})} \tag{4.4}$$

where $P_T$ is the transmitted power in one RB, $G_T$ is the antenna gain of the transmitter, $G_R$ is the antenna gain of the receiver, $L_{j,k}$ is the path loss at distance $d_{j,k}$, and $P_n$ is the power of the additive white Gaussian noise (AWGN). $I_{j,k}$ is the interference power received by the interfered receiver of cluster j from the interferer transmitter of cluster k in the worst case positions WP1 and WP2.

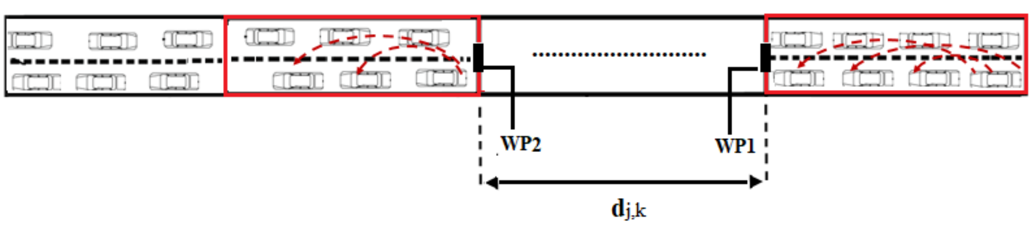


**Figure 4.7:** The worst case positions WP1 and WP2 for sidelink V2V.

Based on the above conditions, the mode selection criterion in Algorithm 4.2 is as follows: If the average values of SINR for both uplink and downlink in the j-th cluster are higher than the average value of SINR for SL (i.e. min $(\bar{\gamma}_j^{UL}, \bar{\gamma}_j^{DL}) > \bar{\gamma}_j^{SL}$ ), the algorithm will move on to check that there are sufficient physical resources to be used for serving the V2V links in cellular mode before taking the decision for switching(lines 5,6). If one of the conditions is not satisfied, the cluster will stay in sidelink mode ($\alpha_j$=1) (lines 11,14). Besides, when there are no more available RBs left for data transmissions, the algorithm allows users of cluster j to reuse the resources of V2V links within

**Algorithm 4.2 : MS-RBRS Algorithm**

1. **Inputs:**
   $N_x$ : Number of RBs in the link $x \in$ {UL,DL}..
   C: number of clusters.
   V(j): vehicles in cluster j,
2. **Initialization**: $\rho^{x}(t)=0$
3. **For** each cluster $j \in C$
4. Compute the average value of $\bar{\gamma}_j^{UL}, \bar{\gamma}_j^{DL}$ , $\bar{\gamma}_j^{SL}$ among all vehicles within cluster j using eqs. (4.1) and (4.2);
5. **If** $Min\left(\bar{\gamma}_j^{UL}, \bar{\gamma}_j^{DL}\right) \geq \bar{\gamma}_j^{SL}$
6. **If** $\Gamma_j^{x} \leq N_x - \rho^{x}$ {check that the required RBs by cluster j are available in each link $x \in$ [UL, DL]}.
7. $\alpha_j = 0$ cluster j operates in cellular mode
8. $\rho^{UL} = \rho^{UL} + \Gamma_j^{UL}$
9. $\rho^{DL} = \rho^{DL} + \Gamma_j^{DL}$
10. **Else If** $\Gamma_j^{SL} \leq N_{UL} - \rho^{UL}$ { check that the required RBs by cluster j are available in UL
11. $\alpha_j = 1$ {cluster j operates in sidelink mode}
12. $\rho^{SL} = \rho^{SL} + \Gamma_j^{SL}$
13. **Else**
    **{Check for RBs reuse in SL}**
14. $\alpha_j = 1$ { cluster j operates in SL mode}
15. Find the cluster *k* with Max $d_{j,k}$
16. compute $\gamma_{j,k}$ between the worst-case positions in clusters k and j
17. **If** $\gamma_{j,k} \geq \gamma_{th}$
18. RBs can be reused in both clusters in SL
19. **Else**
20. RBs cannot be reused in both clusters in SL
21. **End**
22. **End**
23. **Else**
24. Repeat steps 10 - 22
25. **End**
26. **End**
27. **Output:** $\alpha_j$ , Number of RBs for UL $= \rho^{UL}$ , Number of RBs for SL $= \rho^{SL}$ , $j = 1,..,C$

cluster k, where k refers to the cluster with maximum distance $d_{j,k}$ from cluster j, if the minimum requirement of the SINR (i.e., $\gamma_{j,k} > \gamma_{TH}$.) is met. Otherwise, the system will be in outage.

#### 4.4.2.2 Performance Evaluation for MS-RBRS

Our simulation model is based on two cells respectively configured at two gNBs. Each cell has a channel organized in 80 RBs for UL and 80 RBs for DL. Each RB is composed by 12 subcarriers with subcarrier separation $\Delta f$=30 kHz, which corresponds to one of the 5G NR numerologies defined in [62]. The model considers vehicular UEs communicating through cellular mode (uplink / downlink) and via sidelink (direct V2V). The users move along a 2-lane highway and are assumed to enter the cell coverage following a Poisson process with arrival rate ($\lambda_v$). All relevant system and simulation parameters are summarized in Table 4.2. The presented evaluation results intend to assess and illustrate the performance of the MS-RBRS strategy in terms of latency, packet success rate and RB utilization. As reference for comparison, we assume the MSBS strategy of subsection 4.4.1 that takes into account only the quality of the links and the available resources with no RBs reuse.

Figure 4.8 depicts the throughput delivered in Kbits/sec for V2X service aggregated for sidelink and uplink in each cell. The figure illustrates the behavior of the MS-RBRS strategy and the MSBS scheme. Results are presented for different vehicle arrival rates and packet sizes. Here, we can observe that the MS-RBRS strategy outperforms the MSBS strategy in terms of throughput. The MS-RBRS strategy achieved maximum in uplink when the vehicle arrival rate ($\lambda_v$) is 8 vehicles/s and the packet size is 1200 bytes. For the MSBS scheme, the maximum throughput is only 402 Kb/s in uplink (i.e. MS-RBRS strategy achieves a relative gain of 12 %). The reasons are two-fold. First, when the vehicle arrival rate $\lambda_v$ of V2X UEs is increased, more users will use the network and this will increase the number of V2X packets and request more RBs to be used in transmissions. Second, as the number of V2X packets increases, requiring more radio resources, the proposed solution ensures more RBs which can be used to transmit data by taking advantage from the reuse of RBs for other SL V2V links that meet the interference requirements, while the MSBS strategy provides a lower number of available RBs to be used in data transmissions.

In Figure.4.9, we investigate the probability of having outage (i.e. the probability that there are no sufficient RBs to serve all the transmission requests) at a certain point of time. The outage probability

TABLE 4.2: System Parameters used in the Simulation of MS-RBRS

| Parameter | Values |
|---|---|
| Cell radius | 500m |
| Number of RBs per cell | $N_{RB}^{UL}$ =80 RBs |
| | $N_{RB}^{DL}$ =80 RBs |
| Frequency | 5.9GHz |
| Δf | 30 kHz |
| Path loss model | The path loss and the LOS probability for cellular mode are modeled as in [63]. In sidelink mode, all V2V links are modeled based on freeway case (WINNER+B1) with hexagonal layout [ITU-R] [64]. |
| height of the gNB | 10m |
| Base station receiver noise figure | 9 dB |
| Spectral efficiency model to map SINR. | Model in section A.1 of [65]. The maximum spectral efficiency is 1 b/s/Hz. |
| Shadowing standard deviation | 3 dB in LOS and 4 dB in NLOS. |
| Base station antenna gain | 5 dB |
| Length of the highway | 3Km |
| Number of lanes | 2 in one direction |
| Lane width | 4 m |
| Size of clusters | 12 per cell |
| Size of cluster | 250 |
| Vehicular UE height | 1.5m |
| Transmitted power per RB | 23 dBm |
| UE antenna gain | 3 dB |
| vehicle speed | 80 Km/h |
| UE Noise power Pn | -114 |
| Vehicle arrival rate $\lambda v$ | Varied from 1 to 8 vehicles/s |
| Packet arrival rate $\lambda a$ | 1 packets/s |
| packet size (Sm) | {300, 800 , 1200 }bytes |
| TTI duration (Ts) | 0.5ms |
| Time window T | 3s |
| $\gamma_{TH}$ | 14 dB |

of the MS-RBRS strategy and MSBS strategy are plotted against the V2X packet generation rate $\lambda_v$. we can clearly observe that increasing the traffic load leads to an increase in the outage probability of service. It can be also noted that the MS-RBRS strategy can substantially reduce the probability of having outage (i.e. MS-RBRS reduces the probability of having outage from 0.32 to 0.26 compared to MSSB scheme, when the vehicle arrival rate is 8 vehicles/s and the packet size is 1200

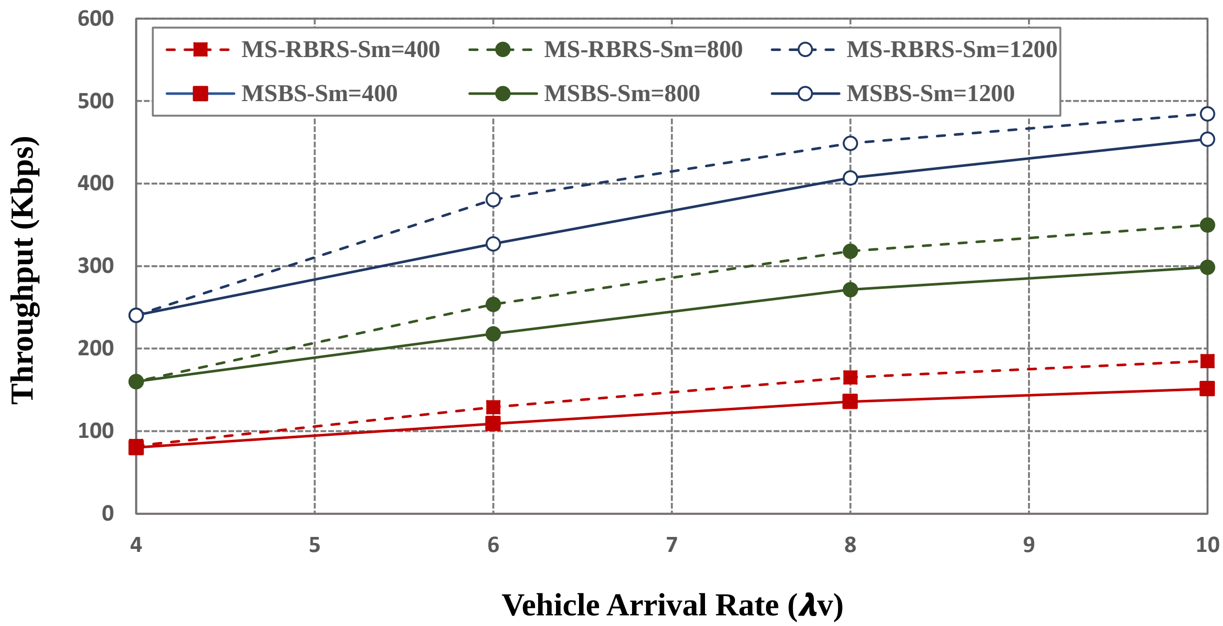


**Figure 4.8:** Aggregated UL and SL throughput as a function of the V2X UE arrival rate $\lambda_v$ (vehicles/s).

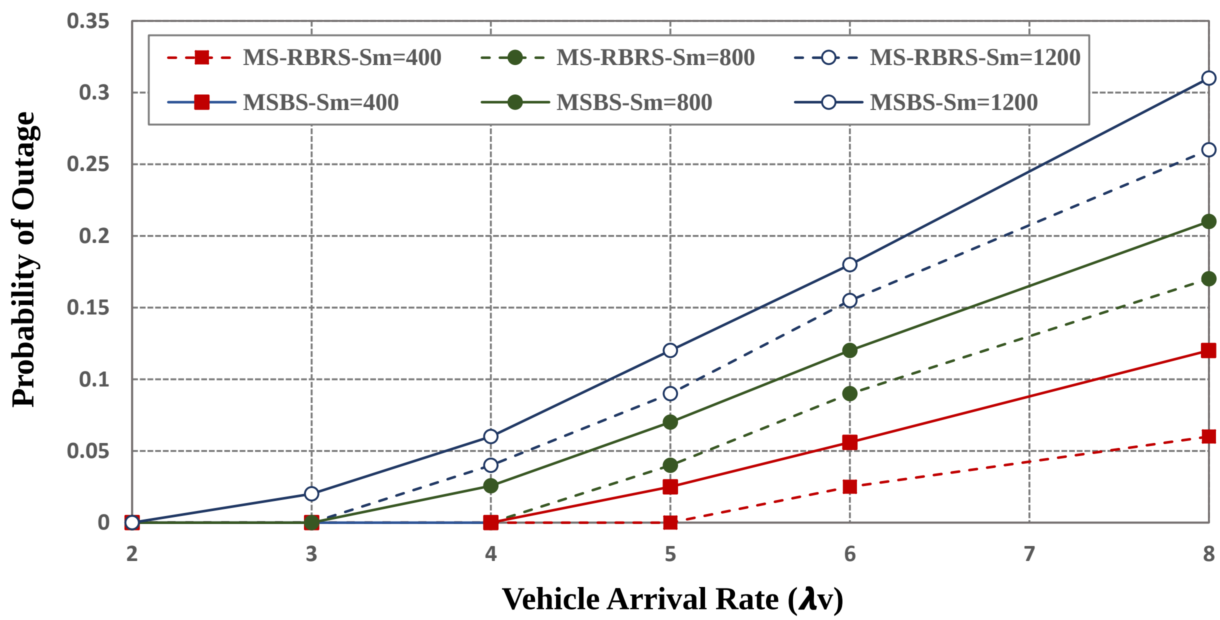


**Figure 4.9:** Outage probability vs V2X UE arrival rate $\lambda_v$ (vehicles/s).

bytes). This is because the proposed solution guarantees more resources to avoid outage.

Figure. 4.10 illustrates the average latency for V2X service caused by channel access delay and the transmission delay. Latency is measured as the time spent by a packet in the system, from the time it is generated until it has been transmitted. We can clearly observe from Figure 4.10 that the average latency increases for all the approaches when vehicle arrival rate ($\lambda_v$) is increased because more vehicles will use the network and request more RBs to be used for the transmissions. This causes an increase in the waiting time and therefore increases the latency. From the presented results, we notice that the proposed solution reduces the latency compared to MSBS strategy. Specially, the average

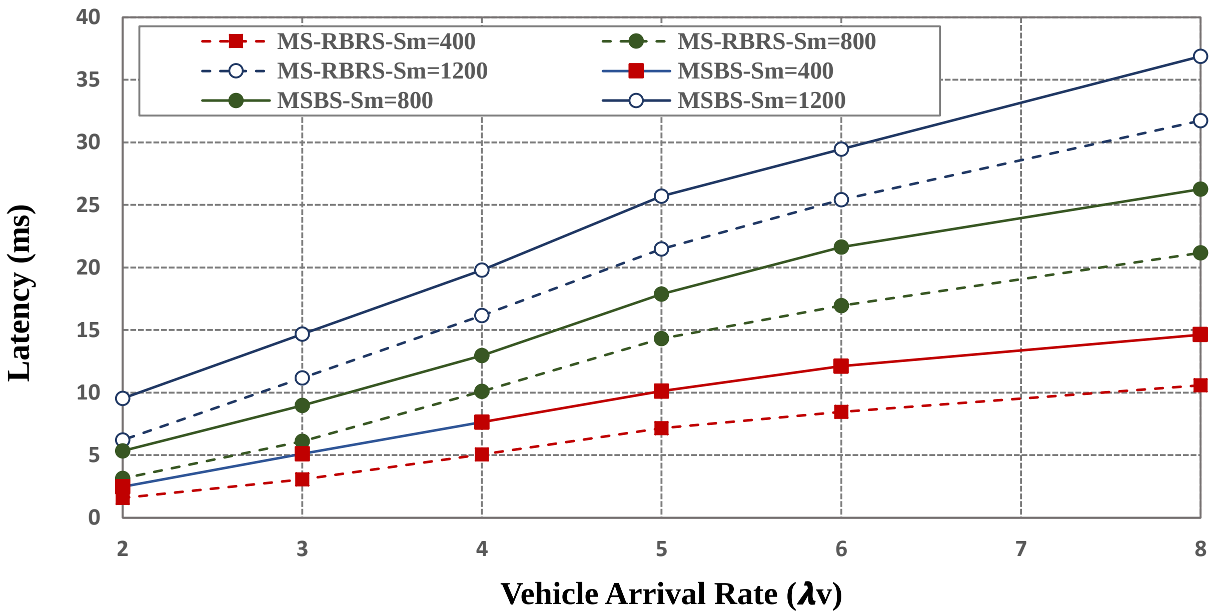


**Figure 4.10:** Average Latency vs V2X UE arrival rate $\lambda_v$ (vehicles/s).

Figure 4.11 presents the RB utilization for V2X service in the UL in each cell as a function of the packet generation rate ($\lambda_v$) for V2X users. The figure shows that there is a marked increase in RBs utilization for all the approaches when vehicle arrival rate ($\lambda_v$) is increased. The figure shows that MS-RBRS strategy maintains the same RBs utilization as the MSSB strategy in different load scenarios. This is because the proposed approach does not allocate additional RBs to serve the traffic and improve the performance in terms of throughput, probability of outage, and latency, but reuses other radio resources used by other SL V2V links that meet the interference requirements.

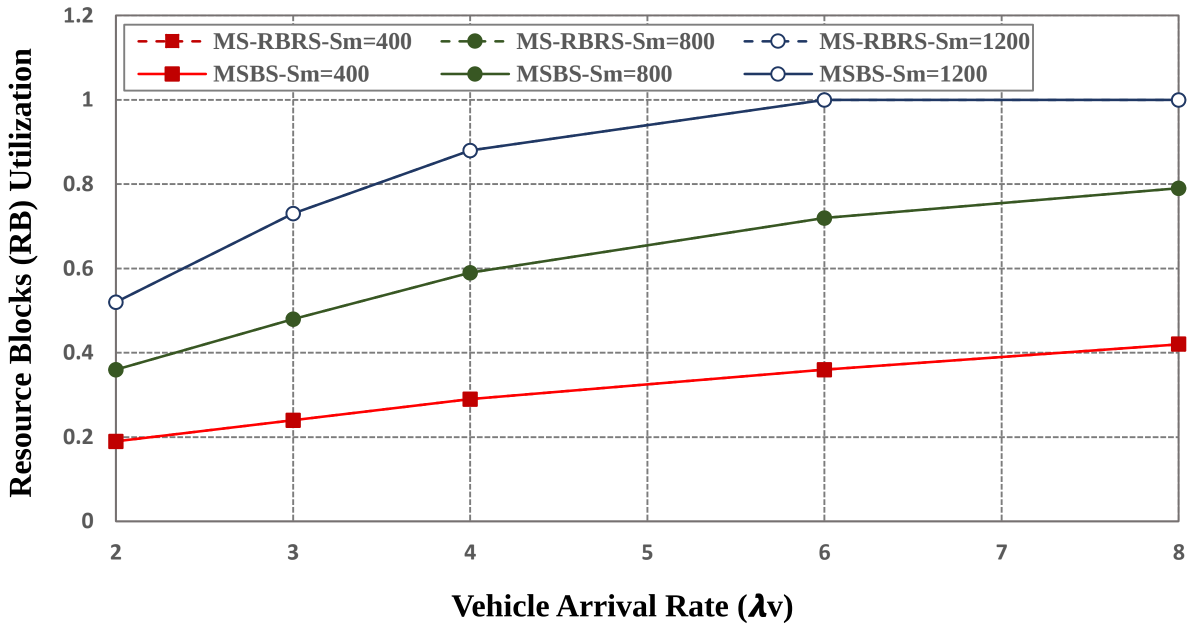


**Figure 4.11:** Uplink RB utilization as a function of the V2X UE arrival rate $\lambda_v$ (vehicles/s).

## 4.5 Concluding Remarks

In this chapter, we have highlighted the main challenges related to the mode selection problem of V2V communication. We have considered different scenarios in which the V2V communications can be performed using: i) direct V2V communication that allows a UE at a vehicle to communicate directly to other UEs in close proximity over the PC5 interface, and ii) cellular mode uses the Uu interface with a two-hop transmission via a base station, i.e. involving UL and DL transmission. In order to face different challenges in terms of the use of radio resources, system throughput and the latency while guaranteeing the signal-to noise-and-interference ratio in both sidelink and cellular links, we have proposed two novel mode selection strategies. In the first strategy, we consider a mode selection algorithm takes into account the quality of the links between V2V users in sidelink mode and between the base station and the vehicles in cellular mode, and the available resources. In addition, in order to achieve further improvements in the use of resources, latency, and congestion, another novel mode selection strategy that exploits resource reuse is introduced, where multiple V2V links may share the same radio resources. The performance of the proposed approaches has been extensively studied via simulations to demonstrate the capability to improve the system performance for V2V communication in terms of RB utilization, throughput, latency of V2X services, and probability of outage. Our simulation results have shown that substantial gains can be achieved with the proposed mode selection strategies. The conducted simulations have revealed first that the MSBS strategy is beneficial in terms of radio resource utilization and congestion comparing to the reference scheme in which all UEs work in sidelink mode. More specifically, when the vehicle arrival rate ($\lambda_v$) is 5 vehicles/s and the packet size is 1200 bytes, the MSBS strategy offers gains of up to 28 % in terms of Resource Block utilization. Additionally, the proposed approach reduces the probability of outage in relation to the reference scheme by around 0.12 % when the vehicle arrival rate ($\lambda_v$) is 6 vehicles/s and the packet size is 1200 bytes.

From the presented results, we also notice that further improvements are obtained by the MS-RBRS strategy in terms of the throughput (i.e. MS-RBRS strategy achieves a relative gain of 12 % compared to the MSBS strategy). Moreover, results have also demonstrated an improvement in the latency (i.e. MS-RBRS strategy achieves a relative gain of up to 17 % compared to the MSBS when the vehicle arrival rate ($\lambda_v$) is 8 vehicles/s and the packet size is 1200 bytes) and in the probability of having outage (i.e. the MS-RBRS strategy achieves a relative gain of up to 9 % compared to the MSBS when the vehicle arrival rate ($\lambda_v$) is 8 vehicles/s and the packet size is 1200 bytes).

# Chapter 5

# Joint Mode Selection and Radio Resource Allocation Strategy for V2X Services in 5G Networks

## 5.1 Introduction

V2V communication is envisioned as one of the most promising enablers for intelligent transportation systems (ITSs) [13]. V2X safety services aim at reducing the risk of traffic accidents. In this regard, ETSI has standardized two safety messages: CAM and DENM [31, 32]. V2X service is extremely sensitive to latency, because the latency is directly related to driving safety.

According to [36], a wide range of V2X use cases such as cooperative collision avoidance and cooperative sensing require the exchange of planned trajectories in 10 ms with very high reliability (>99%) especially, in urban environments. As discussed in chapter 2, LTE-V2X is introduced to meet the requirements of V2X [12]. In turn, 5G V2X, is introduced to address more advanced use cases such as Cooperative collision avoidance, lane merging, and platooning, which have more stringent requirements [14]. NR V2X provides further enhancements to support the stringent latency and reliability requirements for V2X use cases by supporting mini-interval scheduling, where UEs can start their transmissions at any of the 14 OFDM symbols and can occupy any number of OFDM symbols within the slot [14]. In addition to the benefits of mode selection discussed in chapter1, an efficient radio resource scheduling for V2V systems that takes advantage of both centralized and distributed scheduling strategies will be beneficial in order to face various challenges and achieve further improvements in terms of capacity, latency, reliability, and scalability caused by the increase in network size and demands for radio resources.

According to current LTE standards, V2V services can be supported in the SL through two different options, V2V- centralized scheduling and V2V-distributed scheduling [21]. In centralized scheduling, the base station (i.e., the eNB in case of LTE or the gNB in case of future 5G NR) selects and schedules the exact resources to be used by the different sidelink transmissions. In distributed scheduling, UEs autonomously select the radio resources for sidelink transmission from a specific pool of resources indicated by the base station. For this purpose, UEs use a sensing-based semi-persistent scheduling (SPS). However, selecting the appropriate operating mode and the radio resource allocation strategy according to the requirements of V2V communications becomes a challenging issue. In this respect, we propose a novel solution for the resource allocation problem including jointly mode selection and radio resource scheduling with the objective of minimizing the probability of exceeding the maximum delay under the constraint of satisfying the reliability requirement for the V2V communications.

Regarding the mode selection, we propose a novel mode selection strategy to select the appropriate mode of operation. Different from the MSBS strategy proposed in subsection 4.4, in addition to the quality of the links between V2V users in sidelink mode and between the base station and the vehicles in cellular mode, and the available resources, the mode selection proposed in this chapter taking into account the end to end latency. Moreover, the proposed mode selection strategy also decides the appropriate radio resource allocation scheme for the V2V communications. As for the resource allocation, we propose a centralized and a distributed scheduling strategy for cellular and sidelink modes, respectively.

The rest of the chapter is organized as follows. Section 5.2 covers the state of art and contributions. In Section 5.3, we present the system model and study the resource allocation problem including jointly mode selection and radio resource scheduling. To solve the problem, section 5.4 provides a solution for mode selection and radio resource allocation. The proposed solution composes three novel algorithms for mode selection, centralized resource allocation, and distributed resource allocation for V2X networks. Section 5.5 presents the performance evaluation followed by the conclusions in Section 5.6.

## 5.2 State of the Art and contributions

In the recent literature, there have been different studies carried out to investigate the mode selection and radio resource allocation for vehicular communications, as discussed in the following.

Different device-to-device (D2D) communication mode selection and resource optimization algorithms are proposed in [66, 83-85] taking into account the quality of the links for both cellular and D2D links and the interference situation when sharing the radio resources. In [86], a dynamic mode selection and subchannel allocation for an orthogonal frequency-division multiple access (OFDMA) cellular network with D2D communications is introduced to minimize the average end-to-end delay performance under the dropping probability constraint. A mode selection and resource optimization strategy based on channel probability statistical characteristics in the 5G communication network is proposed in [87] to maximize the total throughput of the system and reduce the communication interference between the users.

In relation to radio resource allocation for V2V communications, different works proposed solutions for the centralized approach, in which a central controller within the base station makes decisions for each vehicle based on, e.g. channel state information (CSI). In [88], a heuristic approach has been proposed to solve the optimization problem of the radio resource management for D2D -based V2V communication by converting the requirements for reliability and latency of vehicle communications to optimization constraints. In [89], a resource allocation and spectrum sharing scheme has been designed according to various requirements for different types of links, i.e., high capacity for vehicle to-infrastructure (V2I) links and ultra-reliability for V2V links. Some of the existing solutions for centralized approaches are based on the slowly varying large-scale fading information [88] and the sum V2I ergodic capacity is optimized with V2V reliability guaranteed [89]. However, the main challenge of these approaches is the need for the D2D CSI at the base station level which suffers from a trade-off between the large amount of overhead to collect global information (i.e. especially in scenarios where the channels vary rapidly with time) and the imperfect knowledge of the channels states [22] - [25].

Reference [90] proposed an algorithm that uses graph partitioning tools to divide highly interfering V2V links into different clusters before formulating the spectrum sharing problem in order to maximize the sum V2I capacity while guaranteeing the reliability of all V2V links. The radio resource management (RRM) for V2X services based on D2D communication where both V2V and

V2I communications coexist is investigated in [91]. They proposed an efficient resource allocation and power control for V2X communications algorithm that aims at maximizing the sum rate for V2I users and at guaranteeing the reliability and latency requirements for V2V users. In order to guarantee the latency and reliability requirements of vehicular users while maximizing the information rate of other cellular users, [92] proposed optimal schemes for the radio resource allocation, including power allocation, and modulation/coding selection for the V2V communications. In [93], the authors proposed an algorithm focused on network-controlled resource management based on the knowledge of the position of vehicles. A novel centralized time-division multiple access (TDMA)-based scheduling protocol for vehicular networks is proposed in [94]. They considered a roadside unit (RSU), as a central control unit that makes scheduling decisions based on collected individual information and the channel state information of the communication links within its communication coverage.

Reference [95] proposed a resource allocation scheme based on a network controlled direct V2V communication to adapt to the traffic model and service requirements of V2V. A Location-based Centralized Scheduling Algorithm (LB-CSA) and a Location-based Distributed Scheduling Algorithm (LB-DSA) for V2V broadcast services were proposed in [96] to improve resource utilization efficiency, delay, and controlling packet error rate at an acceptable level. An architecture design for a vehicular network involving both communications through cellular V2X and IEEE 802.11p and where the V2V links are controlled by the cellular eNB is proposed in [97]. In their proposed scheme, every vehicle periodically checks its packet lifetime and requests the cellular eNB to determine V2V links and allocate suitable channels to minimize the total latency. Reference [28 98] presents a graph-based resource allocation scheme for sidelink broadcast V2V communications, where vehicles and spectrum resources are represented by vertices whereas the edges represent the achievable rate in each resource based on the signal–to–interference–plus–noise ratio (SINR) that vehicles perceive.

Some other research studies have been carried out in the context of decentralized resource allocation for V2V communications making use of different approaches such as matching game [99], reinforcement learning [100], or location-partition [101], where the most critical challenges are limited bandwidth, reliability and latency. Specifically, a novel proximity and quality-of service-aware resource allocation framework for V2V communication is proposed in [99]. The proposed scheme incorporates the physical proximity and traffic demands of vehicles where V2V links are clustered based on the positions of the vehicles and traffic load to minimize queuing latency and satisfy reliability requirement. A low-complexity outage-optimal distributed resource allocation mechanism for V2V communications based on deep reinforcement learning is introduced in [100]. In their proposed approach, each V2V link is supported by an autonomous agent, which makes decisions to find the optimal sub-band and power level for transmission without requiring or having to wait for global information to satisfy the stringent latency constraints on V2V links while minimizing the interference to V2I communications. A novel location-partition-based resource allocation scheme is proposed in [101] to improve the minimum achievable rate and satisfying the requirement that the interference generated by all V2V links is below a certain threshold.

From an analysis perspective, [102] presented a detailed analysis of the performance of LTE-V sidelink Mode 4, and introduced a modification to its distributed scheduling. The authors in [103] focus on autonomous V2V resource reselection schemes, which allow communications also in out-of-coverage areas. Similarly, [104] introduced a new approach to radio resource allocation for V2V in case of out of-coverage areas that are delimited by network infrastructure based on collecting and predicting information such as vehicle velocity, density and message traffic, used by the network infrastructure to ensure the reliability of the V2V services. To deal with the new challenges in V2V communications, a decentralized resource allocation scheme has been designed by [100] that benefits from the users' knowledge of their local D2D channel state to intelligently manage the spectrum access.

Despite the existence of the abovementioned works in the area of radio resource management, none of the above works has considered jointly the mode selection between sidelink and cellular modes and the radio resource scheduling taking into account the following three aspects together: the quality of the links between V2V users in sidelink mode and between the base station and the vehicles in cellular mode, the available resources, and the end to end latency. This constitutes a novelty of this chapter. Another novelty of this chapter relies on the considered radio resource allocation solution using two radio resource allocation approaches: distributed radio resource allocation for sidelink mode and centralized radio resource allocation to achieve low latency allowing for transmission over a fraction of a slot, referred to as "mini-slot" transmission.

Based on all the above considerations, the key contributions of this work are summarized as follows:

- We formulate the resource allocation problem for V2V communications including mode selection and radio resource scheduling with the objective of minimizing the probability of exceeding the maximum delay under the constraint of satisfying the reliability requirement.

- To solve the resulting problem, we propose a novel mode selection to decide when it is appropriate to select sidelink mode and use a distributed approach for radio resource allocation or cellular mode and use a centralized radio resource allocation. The proposed mode selection strategy takes into account the quality of the links between V2V users in sidelink mode and between the base station and the vehicles in cellular mode, the available resources, and the latency.

- A novel low-latency resource allocation strategy is proposed for V2V communications using two radio resource allocation approaches: distributed radio resource allocation for sidelink mode and centralized radio resource allocation for cellular mode. The proposed strategy supports dynamic assignments by allowing transmission over a fraction of a slot (i.e. mini-slot) in order to support a more efficient approach to achieve low latency and improve the system performance in terms of resource utilization and achievable data rate.

- The performance of the proposed approach is evaluated using extensive simulations to demonstrate its capability to perform an efficient allocation of resources among users in terms of latency, packet success rate and resource utilization. Simulation results show the effectiveness of the proposed schemes under different system parameters.

## 5.3 System Model and Problem Formulation

### 5.3.1 System Model

We consider a cellular network that consists of a straight highway covered by a gNB with one cell. The model assumes several independent vehicles moving along the highway divided into clusters (i.e., smaller zones according to the length of the road), as illustrated in Figure 4.1. We assume an indicator $\alpha_j$ to reflect the operation mode of the j-th cluster (i.e., $\alpha_j$ =1 if vehicles within the cluster operate in sidelink mode, and $\alpha_j$ =0 if vehicles within the cluster operate in cellular mode). The vehicles in the highway are assumed to enter the cell coverage following a Poisson process with arrival rate $\lambda_v$ and move with speed of v km/h. Each vehicle is assumed to generate packets randomly with rate $\lambda_a$ according to Poisson arrival model. The length of the packets is $S_m$. Each UE transmitter has a queue buffer to store packets waiting to be transmitted. In practice, we assume infinite maximum capacity for the queue and the packets move out of the queue only when they are transmitted to other vehicles. The notations used throughout this work are listed in Table 5.1.

### 5.3.2 Radio Resource Model

We assume that the cell operates with 5G NR technology using OFDMA access configured with a number of radio resources that are allocated in the time/frequency domains for uplink and downlink. In the frequency dimension, the cell bandwidth is divided into Resource Blocks (RBs) each one consisting of 12 contiguous subcarriers with subcarrier separation Δf corresponding to one of the numerologies defined in [62]. The time domain is organized in time slots of 14 symbols, each one with duration $T_s = 1/\Delta f + T_{CP}$, where $T_{CP}$ is the cyclic prefix duration.

The RB is the smallest resource allocation unit that we consider in our model and, in the time domain, it can be assigned in mini-slots to provide low latency payloads with an immediate start time without needing to wait for the start of a slot boundary. A mini-slot is defined as a number of consecutive symbols (i.e., 2, 4, or 7 symbols) [105]. The number of available RBs in downlink is denoted as $N_{DL}$. As for the uplink and the sidelink, and given that the use of sidelink in 5G has not been yet standardized in the current release 15, but it is still a subject under study in release 16 [15], we assume that, like in LTE, the sidelink RBs are a subset of the uplink RBs. Therefore, the total RBs of the uplink $N_{UL}$ are divided into two subsets, namely $N_{RB}^{UL}$, which corresponds to the set of UL RBs that are used for uplink transmissions between the UEs and the gNB, and $N_{RB}^{SL}$ which corresponds to the set of UL RBs used to support sidelink transmissions. To model the resource allocation, we assume a binary variable denoted by $u_{j,i,k,t}^{x} \in \{0, 1\}$, which indicates the assignment of the *k*-th RB of the set $x \in \{SL,UL,DL\}$, to the *i*-th UE in cluster *j* at time *t* (i.e., , $u_{j,i,k,t}^{x}$ =1 whenever whenever the UE of vehicle *i* is transmitting over RB *k* from the radio resource pool of $x \in \{SL,UL,DL\}$, and $u_{j,i,k,t}^{x}$

$u^{x}_{j,i,k,t}$ =0 otherwise).

TABLE 5.1: Summary of notations.

| Notation | Definition |
|---|---|
| $N_{UL}$ , $N_{DL}$ | Set of UL, and DL RBs. |
| $u^{x}_{j,i,k,t}$ | Indicates the assignment of the *k*-th RB of the set x∈{SL,UL,DL}, to the *i*-th UE in cluster *j* at time *t* |
| $m(j,i,t)$ | The number of transmitted packets by the i-th vehicle of the *j*-th cluster at each time *t* within the time window T |
| $SP_{eff,x}$ | The spectral efficiency associated to the modulation and coding scheme to be used in the *x* link, |
| $T_s$ | The symbol duration. |
| $W^{SL}_{j,i}(l)$ , $W^{CL}_{j,i}(l)$ | The latency of packet *l* of user *i* when the cluster *j* operates in sidelink (SL) and in cellular (CL) mode. |
| $W^{SL}_{T,j,i}(l)$, $W^{SL}_{Q,j,i}(l)$ | The transmission delay and queueing delay the of the *l* -th packet from UE *i* in the SL buffer. |
| $W^{UL}_{Q,j,i}(l)$ , $W^{DL}_{Q,j,i}(l)$ | The queueing delay the of the *l* -th packet from UE *i* in the UL and DL buffers. |
| $W^{UL}_{T,j,i}(l)$ , $W^{DL}_{T,j,i}(l)$ | The UL transmission delay and DL transmission delay for the *i*-th UE of the *j*-th cluster. |
| $\alpha_j$ | Indicator $\alpha_j$ to reflect the operation mode of the *j*-th cluster |
| $N^{UL}_{T,i}$ , $N^{DL}_{T,i}$ | The total number of packets that have been transmitted in uplink and downlink by user *i* |
| $\rho^{x}$ | The total number of RBs that have been allocated to link $x$∈{SL,UL,DL}. |
| $\Gamma^{x}_{j}$ | The average number of required RBs from V2X users in cluster *j* in $x$∈{SL,UL,DL}over the time window *T*. |
| $D(i,l)$ | The remaining time for each packet *l* generated by vehicle *i* |
| $\beta_x(i,l)$ | The resources requested by user *i* to transmit packet *l* in link x∈{SL,UL,DL} |
| $D^{x}_{TTI}(i,l)$ | The size of TTI in link $x$∈{SL,UL,DL} for transmitting the *l*-th packet of the *i*-th user |
| $W_{th}$ | The maximum acceptable latency |
| $T_g(i,l)$ | The time when the packet l of user i was generated. |

### 5.3.3 V2V Latency Model

In V2V communications, latency is considered as one of the most critical requirements rather than data rate. In our considered V2V scenario, the latency is mainly due to two types of delays, namely the queueing delay, and the transmission delay. The queuing delay is defined as the time that the *l*-th packet of the *i*-th UE waits in a queue until the time when the transmission starts. The queuing delay depends on different aspects such as the number of users that are multiplexed on the radio resources, the amount of generated packets, the size of packets, etc. The transmission delay is defined as the number of symbols (i.e., 2, 4, or 7 symbols) needed to transmit the packet and depends on the available resources, packet length, radio channel conditions, etc.

The total latency of the *l*-th packet of vehicle *i* in cluster *j* is defined as

$$W_{j,i}(l) = \begin{cases} W_{j,i}^{SL}(l) & \alpha_j = 1 \\ W_{j,i}^{CL}(l) & \alpha_j = 0 \end{cases} \tag{5.1}$$

Where $W_{j,i}^{SL}(l)$ and $W_{j,i}^{CL}(l)$ denote the latency when the cluster operates in sidelink and in cellular mode, respectively. The computation of these terms is detailed in the following sub-sections.

#### 5.3.3.1 Latency Model in the Sidelink Mode

The SL latency $W_{j,i}^{SL}(l)$ is the time duration required for the vehicle *i* to transmit a packet *l* to other vehicles of its cluster *j* through the SL. It accounts for the queuing delay of the *l*-th packet in the buffer and the transmission delay, denoted respectively as $W_{Q,j,i}^{SL}(l)$, and $W_{T,j,i}^{SL}(l)$. Then:

$$W_{j,i}^{SL}(l) = W_{Q,j,i}^{SL}(l) + W_{T,j,i}^{SL}(l) \tag{5.2}$$

where $W_{Q,j,i}^{SL}(l)$ includes the time packet takes until the transmission is scheduled in SL, and the time from the moment that packet is scheduled to the beginning of the TTI decided for the SL transmission. The transmission time $W_{T,j,i}^{SL}$ is given by $W_{T,j,i}^{SL} = S_m / R_{j,i}^{SL}$, where $R_{j,i}^{SL}$ is the instantaneous data rate of user *i* within cluster *j* in sidelink.

#### 5.3.3.2 Latency Model in the Cellular Mode

In UL/DL-based V2V operation, each UE *i* in a cluster transmits the packets to the gNB through the uplink. Then, the gNB multicasts every packet to the other UEs of the cluster over the DL. Therefore, the latency of the $l$-th packet from UE *i* in cellular (CL) mode ,$W_{j,i}^{CL}(l)$ considers both the UL transmission latency $W_{j,i}^{UL}(l)$ and the DL transmission latency $W_{j,i}^{DL}(l)$.

Each UE *i* maintains a queue for the uplink cellular transmissions. On the other hand, we assume that the gNB has different queues for each UE in the downlink. In this case, the transmitting UE should store the packet in the uplink queue until it can be sent to the gNB. Then, the gNB stores the packet it received from the UE in its downlink queue until it can multicast it to the other UEs. Then, let denote as $W_{Q,j,i}^{UL}(l)$, $W_{Q,j,i}^{DL}(l)$, $W_{T,j,i}^{UL}(l)$, and $W_{T,j,i}^{DL}(l)$ the UL queuing delay, the DL queuing delay, UL transmission delay and DL transmission delay, respectively, for the *i*-th UE of the *j*-th cluster. The end to end delay of each packet denoted as $W_{j,i}^{CL}(l)$ , is defined as

$$W_{j,i}^{CL}(l) = \left( \underbrace{\left( W_{Q,j,i}^{UL}(l) + W_{T,j,i}^{UL}(l) \right)}_{W_{j,i}^{UL}(l)} + \underbrace{\left( W_{Q,j,i}^{DL}(l) + W_{T,j,i}^{DL}(l) \right)}_{W_{j,i}^{DL}(l)} \right) \tag{5.3}$$

The transmission time $W_{T,j,i}^{x}$ for each link x ∈ {UL, DL} is given by $W_{T,j,i}^{x} = S_m / R_{j,i}^{x}$. $R_{j,i}^{x}$ is the instantaneous data rate for the transmission in the link $x \in$ {UL,DL}.

### 5.3.4 Problem Formulation

The focus of this chapter to optimize the mode selection and resource allocation mechanisms in order to minimize the probability of exceeding the maximum delay under the constraint of satisfying the reliability requirement for V2V users. For this purpose, we intend to determine on the one hand the operation mode $\alpha_j$, i.e. cellular or sidelink, for each cluster $j \in C$ and, on the other hand, to determine the resource allocation $u_{j,i,k,t}^{x}$ to the different UEs in the cluster.

Based on the above, the optimization problem for the resource allocation is defined as follows:

$$\min_{\alpha_j,\, u_{i,k,t}} \Pr\left(W_{j,i}(l) > W_{th}\right) \tag{5.4}$$

$$\gamma(\alpha_j) \geq \gamma_{\text{TH}} \qquad \text{j= 1,....C} \tag{5.4.a}$$

$$\sum_k u^x_{j,i,k,t} \leq 1, \qquad i \in V(j) \tag{5.4.b}$$

$$u^x_{j,i,k,t} \in \{0,1\}, \quad i \in V(j), \quad k \in K \tag{5.4.c}$$

where $W_{th}$ is the maximum acceptable latency. The reliability constraint is expressed in (5.4.a) and intends to ensure that the average SINR for V2V links of *j*-th cluster in the $\alpha_j$ mode, $\gamma(\alpha_j)$ is larger than a predetermined threshold, $\gamma_{\text{TH}}$. The average SINRs for V2V links of *j*-th cluster in the $\alpha_j$ mode is defined as

$$\gamma(\alpha_j) = \begin{cases} \min\left(\bar{\gamma}_j^{UL}, \bar{\gamma}_j^{DL}\right) & \alpha_j = 0 \\ \bar{\gamma}_j^{SL} & \alpha_j = 1 \end{cases} \tag{5.5}$$

Then, constraint (5.4a) reflects that both average uplink and downlink SINRs, denoted respectively as $\bar{\gamma}_j^{UL}$ and $\bar{\gamma}_j^{DL}$, should be larger than $\gamma_{\text{TH}}$ when the cluster *j* operates in cellular mode, and the average SINRs of sidelink, denoted as $\bar{\gamma}_j^{SL}$, should be larger than $\gamma_{\text{TH}}$ when cluster *j* operates in sidelink mode). The orthogonality of the resource allocation is represented by constraints (5.4.b) and (5.4.c), which means that sub-channel *k* is allocated to only user *i* at time *t*.

## 5.4 Mode Selection and Radio Resource Allocation Strategy (MS-RRAS)

The problem in (5.4) with their constraints is a nonlinear optimization problem. Such an optimization problem is generally hard to solve. Therefore, we propose to split the problem into two parts, as depicted in Figure 5.1. In the first part, we consider a mode selection strategy performed by a controller within the gNB responsible for determining the optimal mode of operation for each cluster $j \in C$ Then, the second part is the radio resource allocation algorithm that performs the process of assigning RBs to the users based on the mode selected by the mode selection strategy of part 1. This runs as a centralized RB allocation at the gNB for the clusters in cellular mode and as a distributed RB allocation running at the users for the clusters in sidelink mode. A detailed description of both the dynamic mode selection and the low complexity heuristic centralized and distributed radio resource allocation solutions is given in sub-sections 5.4.1, and 5.4.2, respectively.

### 5.4.1 The proposed Mode Selection Algorithm

This chapter considers a mode selection approach, in which the controller keeps track of the performance of both modes, i.e. SL and CL, for each cluster in terms of the received signal, availability of resources and latency. Based on this, the algorithm decides the appropriate mode for each cluster and applies it in the real network. In order to assess the performance of the current mode of operation $\alpha_j$ in cluster $j$, the mode selection algorithm retrieves information from the environment This information is denoted as INF($\alpha_j$) in Figure 5.1, and includes the network characterization in terms of traffic generation, vehicle's speed, required amount of resources for V2V links, the SINR received by each vehicle in the cluster and latency measurements of the different packets. Similarly, the mode selection also retrieves information from the RB allocation algorithm, denoted as INF2 in Figure 5.1. This includes the available resources for each link $x \in \{UL,DL,SL\}$ and radio channel conditions. In turn, in order to assess the performance of the mode of operation that is not being used (e.g. to assess the performance of CL mode when operating in SL mode, or the opposite), the mode selection algorithm retrieves information from a network model that simulates the behavior of the other mode in offline manner. This information is denoted as INF($\overline{\alpha}_j$) in Figure 2, where ($\overline{\alpha}_j = 1 - \alpha_j$) denotes the opposite mode of operation to $\alpha_j$. The network model is established to provide an accurate assessment and understanding of the system behavior. It is based on a network simulation that accounts for the road traffic, vehicle's speed, position, traffic generation, etc., according to a given real network characteristic. INF1 refers

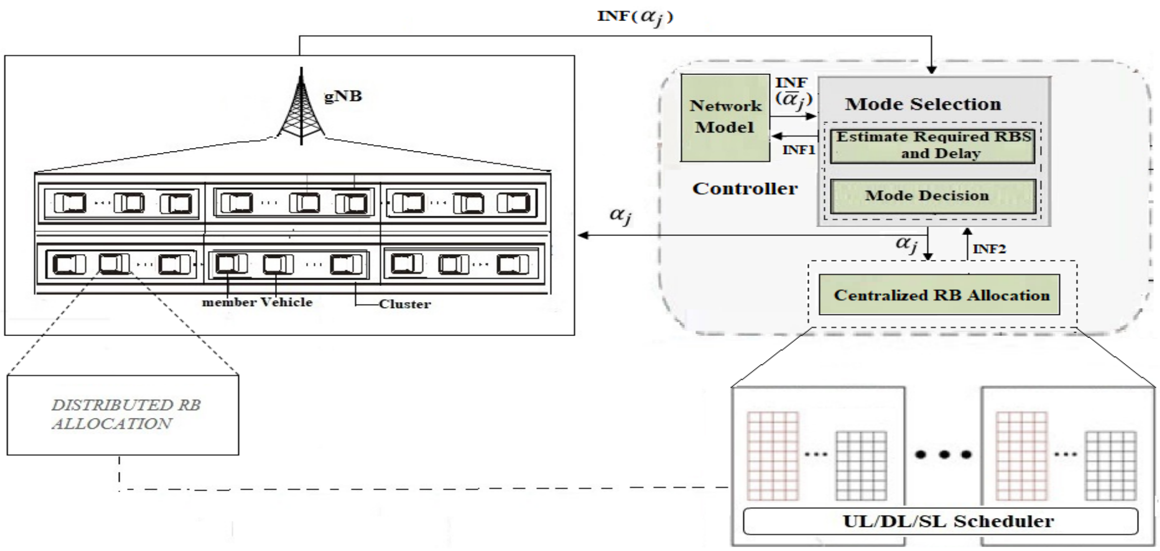


**Figure 5.1:** Mode Selection and Radio Resource Allocation strategy.

to the configuration parameters of the network model. Specifically, INF1 provides information about $\bar{\alpha}_j \in \{0, 1\}$, number of clusters, size of each cluster, number of vehicles with each cluster, traffic generation, etc. The network model uses these parameters as input and generates the corresponding information about the performance of the mode $\bar{\alpha}_j$ as output, namely the required amount of resources for V2V links, the SINR received by each vehicle *i*-th in the cluster, and the latency measurements.

The mode selection algorithm operates periodically in time windows of duration *T* and is detailed in the pseudo-code of Algorithm 5.1. The proposed strategy assumes that cluster *j* is working in mode $\alpha_j$ at a given time window, and the algorithm determines the mode for the next time window. To make this decision, the proposed mode selection strategy considers the following conditions for cluster *j*:

- The cellular V2V links to be established must satisfy the SINR conditions. Specially, it must ensure that both uplink and downlink SINRs should be larger than the SL SINR, if the cluster j-th needs to be switched and operated in CL mode. The average SINR for all the UEs in sidelink, uplink, and downlink transmissions within each cluster in a time window T can be statistically estimated based on (4.1) and (4.2) of Chapter 4.

- The V2V links to be established in sidelink or cellular modes must ensure that the amount of required RBs by cluster j is less than the number of RBs available for the mode in which the vehicles are to be switched and operated (i.e., eq. (4.3) of Chapter 4) must be satisfied (lines 13):

- The V2V links to be established in sidelink or cellular mode should satisfy the latency requirements $W_{TH}$ . For this purpose, the average latencies in cellular and sidelink modes are computed (lines 7). The average latency for all the packets generated by vehicles within each cluster in cellular mode (UL and DL) denoted as $W_j^{CL}$ can be estimated as follows:

$$W_j^{CL} = \left( \frac{1}{V(j)} \sum_{i=1}^{V(j)} \left( \frac{\sum_{l=1}^{N_{T,i}^{UL}} W_{j,i}^{UL}(l)}{N_{T,i}^{UL}} + \frac{\sum_{l=1}^{N_{T,i}^{DL}} W_{j,i}^{DL}(l)}{N_{T,i}^{DL}} \right) \right) \tag{5.6}$$

where $N_{T,i}^{UL}$ and $N_{T,i}^{DL}$ are the total number of packets that have been transmitted in uplink and downlink respectively. Following similar considerations like in the cellular mode, the average latency for all the packets generated by vehicles within each cluster in sidelink mode denoted as $W_j^{SL}$ is given

$$W_j^{SL} = \left( \frac{1}{V(j).N_{T,i}^{SL}} \sum_{i=1}^{V(j)} \sum_{l=1}^{N_{T,i}^{SL}} W_{SL,i}(l) \right) \tag{5.7}$$

Based on the above conditions, the mode selection criterion is as follows: If the average values of SINR for both uplink and downlink in the j-th cluster are higher than the average value of SINR for sidelink, (i.e. min ( $\bar{\gamma}_j^{UL}, \bar{\gamma}_j^{DL}$ )> $\bar{\gamma}_j^{SL}$ ) and value of threshold, $\gamma_{TH}$ , and the number of available

**Algorithm 5.1:** Mode Selection Strategy

1. **Inputs:**
   $N_{UL}$, $N_{DL}$: Number of RBs in UL and DL.
   C: number of clusters.
   V(j): vehicles in cluster j,
   INF($\alpha_j$) , INF($\bar{\alpha}_j$) ,INF1, INF2
2. **Output**: $\alpha_j$ , $j$=1,..,$C$
3. **Initialization**: $\rho^x(t)$ =0
4. **Iteration**
5. **for** each cluster $j \in$ C
6. **for** each vehicle $i \in$ V(j) within cluster j **do**
7. Compute the $W_{j,i}^{CL}(l)$ $W_{j,i}^{SL}(l)$ for all packets generated by vehicle i
8. **End**
9. Compute the average value of $W_j^{CL}$ && $W_j^{SL}$ among all vehicles within cluster j;
10. Compute the average value of $\bar{\gamma}_j^{UL}, \bar{\gamma}_j^{DL}$ , $\bar{\gamma}_j^{SL}$ among all vehicles within cluster j;
11. **If** $Min\left(\bar{\gamma}_j^{UL}, \bar{\gamma}_j^{DL}\right) \geq \bar{\gamma}_j^{SL}$
12. **If** $Min\left(\bar{\gamma}_j^{UL}, \bar{\gamma}_j^{DL}\right) \geq \gamma_{TH}$
13. **If** $\Gamma_j^x \leq N_x - \rho^x$ {check that the required RBs by cluster j are available in each link $x \in$ [UL, DL]}.
14. **If** $W_j^{CL} < W_{TH}$
15. $\alpha_j$ =0 cluster j operates in cellular mode and use centralized radio resource allocation
16. $\rho^{UL} = \rho^{UL} + \Gamma_j^{UL}$
17. $\rho^{DL} = \rho^{DL} + \Gamma_j^{DL}$
18. **Else**
19. $\alpha_j$ =1 cluster j operates in sidelink mode and use centralized radio resource allocation
20. $\rho^{SL} = \rho^{SL} + \Gamma_j^{SL}$
21. **End**
22. **Else**
23. Repeat steps 19-20
24. **End**
25. **Else**
26. Repeat steps 19-20
27. **End**
28. **Else**
29. Repeat steps 19-20
30. **End**
31. **End**

RBs for cellular users in UL and DL is higher than the RBs required by V2V links within the cluster (lines 11-13), the algorithm will move on to check if the average packet latency for all the vehicles within cluster j in cellular mode is less than threshold value $W_{th}$ (line 14). If this is the case, the corresponding vehicles within cluster j will be switched to operate in cellular mode ($\alpha_j$=0) (line15). Otherwise the cluster will stay in sidelink mode ($\alpha_j$=1) (line 19).

### 5.4.2 Radio Resource Allocation Algorithms

In this section, we first present the radio resource allocation algorithm for V2V communications in cellular mode. We then introduce the key steps of the decentralized resource allocation mechanism in sidelink mode.

#### 5.4.2.1 Centralized Radio Resource Allocation Algorithm for Cellular Mode.

In this section, we propose a heuristic centralized resource allocation scheme applicable when algorithm 5.1 selects the cellular mode for a cluster. In this approach, the uplink and downlink allocations are performed dynamically and orthogonal radio resources will be allocated to both links from the sets of $N_{UL}$ and $N_{DL}$ RBs to serve the uplink and downlink queues, respectively. It is assumed that two separate radio resource allocation algorithms ar`e executed for the UL and the DL. For the uplink transmission, thanks to the Buffer Status reporting procedure, UEs can provide the serving gNB with the information about the amount of data available for transmission in their UL buffers to allow the scheduler to determine the amount of resources to grant to each UE.

As reflected by the constraint (5.4b) of the optimisation problem, a given RB in UL or DL can be allocated to, at most one user of the cell in cellular mode. The pseudo-code of the proposed radio resource allocation algorithm is shown in algorithm 5.2. The same algorithm is valid for UL and DL, represented by a generic link $x \in \{UL,DL\}$. The general idea is that, at the scheduling interval (i.e., symbol duration), the algorithm runs sequentially for the number of involved uplink and downlink transmissions in order to assign the RBs that provide the best channel conditions to the users with highest priority (i.e. the user with the more urgent packet). For this purpose, the algorithm calculates the remaining life time for each packet *l* buffered in uplink or downlink queue respectively (lines 6-8) and allocates the best RB in terms of the channel quality to the packet with the with lowest remaining life time from all queues (i.e., the packet *l* with lowest remaining life time $D(i,l)$ will be scheduled first for transmission on the best RB in terms of the channel quality). The remaining life time for packet *l* from the vehicle *i* in cellular mode at time t is calculated as:

$$D(i,l) = \left(W_{th} + T_g(i,l) - t\right) \tag{5.8}$$

where $T_g(i,l)$ is the time when the packet *l* was generated. The proposed scheme supports a flexible structure for dynamic scheduling of users with different TTI sizes using mini-slots of 2, 4 or 7 symbols in accordance to each user requirements. For this purpose, the algorithm first needs to select the value of the TTI size, denoted as $D_{TTI}^{x}(i,l)$, and obtain, $\beta_x(i,l)$ which denotes the resources requested by user *i* to transmit packet *l* and is defined as:

$$\beta_x(i,l) = \frac{S_m}{D_{TTI}^{x}(i,l).\, SP_{eff,x} \cdot B} \tag{5.9}$$

where x denotes the type of link, i.e. x $\in$ {UL,DL}.

The algorithm starts with the $D_{TTI}^{x}(i,l)$ equal to 2 symbols, and then computes the required RBs $\beta_x(i,l)$. If $\beta_x(i,l)$ is higher than the available resources, the algorithm will increase the TTI size until ensuring that all the resources requested by user *i* are met or until reaching the maximum TTI size of 7 symbols (lines 10-20). The scheduled packet is moved out from the corresponding queue and the allocated radio resources in each time slot are subtracted from the available radio resources list (lines 22). This procedure continues until either there is no radio resource left or no packets are pending in the queue.

#### 5.4.2.2 Distributed Radio Resource Allocation Algorithm for Sidelink Mode

In this section, the distributed radio resource allocation scheme used by UEs in sidelink mode presented. The pseudo-code of the proposed distributed radio resource allocation algorithm is

summarized in Algorithm 5.3. In the proposed scheme, each UE autonomously selects the radio

**Algorithm 5.2**: Centralized resource allocation Solution

1. **Inputs:**

$N_{RB}^{UL}$ : Number of RBs in UL

$\alpha_j$ : Mode selected by the Control $\alpha_j \in \{0,1\}$

x ϵ {UL, DL}

2. **Iteration**
3. **for** each cluster $j \in C$
4. **If** $\alpha_j = 0$ { vehicles within cluster j operate in cellular mode
5. **for** each vehicle $i \in$ V(j) within cluster j in cellular
6. **for** each packet $l$ in the queue **do**
7. Compute $D(i,l)$ for each packet generated by vehicle i based on (5.8)
8. **End**
9. **End**
10. **Set** $D_{TTI}^{x}(i,l)$ to 2
11. Compute $\beta_x(i,l)$ based on (5.9)
12. **If** Available RBs within $D_{TTI}^{x}(i,l) \geq \beta_x(i,l)$
13. Allocate RBs with best channel condition within $D_{TTI}^{x}(i,l)$ equal to 2 symbols vehicle $i$ with highest priority $(\text{Min}\, D(i,l))$
14. **Else**
15. **Set** $D_{TTI}^{x}(i,l)$ to 4
16. **If** Available RBs within $D_{TTI}^{x}(i,l) \geq \beta_x(i,l)$
17. Allocate RBs with best channel condition within $D_{TTI}^{x}(i,\ell)$ equal to 4 symbols vehicle $i$ with highest priority $(\text{Min}\, D(i,l))$ . .
18. **Else**

    **Set** $D_{TTI}^{x}(i,l)$ to 7

19. Allocate RBs with best channel condition within $D_{TTI}^{x}(i,l)$ equal to 7 symbols to vehicle $i$ with highest priority $(\text{Min}\, D(i,l))$
20. **End**
21. Transmit packet $l$ from the queue of the i-th user in the j-th cluster
22. Remove packet $l$ from the queue.
23. **End**
24. **End**

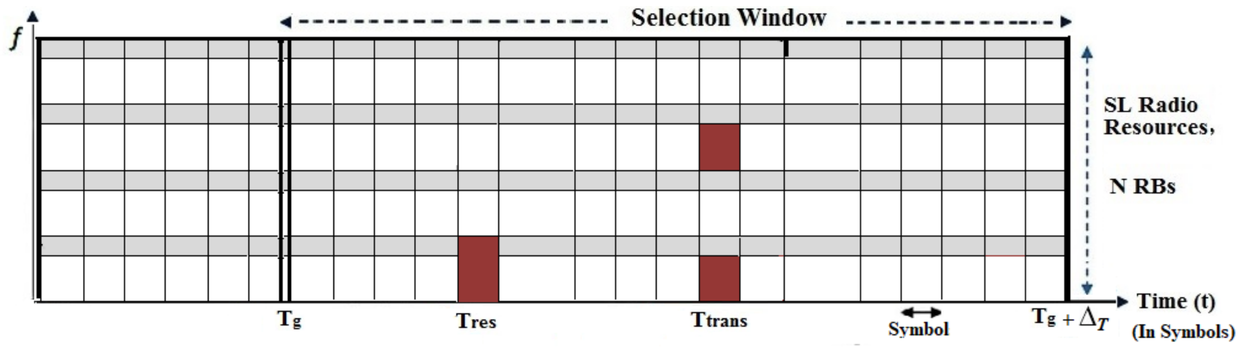


**Figure 5.2**: Resource Selection Strategy.

resources for direct V2V transmission without interaction with the gNB. Like in the autonomous radio resources selection specified in [21, 34,106], the proposed scheme uses the sensing based resource selection (i.e., UE selects unused resources based on a short-term sensing) to determine suitable transmission opportunities, i.e. the number of symbols and RBs required for the transmission. Since there is no central coordinator for the resource scheduling, each UE monitors the resource usage by other UEs before selecting the resource for its V2V transmission. In the proposed strategy, each vehicle *i* generates packet at time $T_g$ and reserves radio resources between $T_g$ and $T_g + \Delta_T$ where $\Delta_T$ is given by $\Delta_T = W_{th} - W^{x}_{T,j,i}(l)$ in order to fulfil the latency threshold. This is called the selection window (Figure. 5.2). In order to effectively reserve radio resources for its transmission and guarantee that the reserved resources are not selected by other UEs to mitigate resource collision, each vehicle *i* needs to send a reservation signal before the packet transmission (line 20). The reservation signal is transmitted at time $T_{res}$ in Figure. 5.2, while the packet is transmitted at time $T_{trans}$. The reservation signal includes the RBs to be reserved, the packet remaining lifetime, the modulation and coding scheme (MCS), and channel conditions. However, for ease of exposition, we focus on the selection of radio resources autonomously to discuss the algorithm in the following.

For cluster $j \in C$, in sidelink mode, each vehicle $i \in V(j)$ needs to autonomously select RBs from the available resources within the selection window and transmit the reservation signal. Based on the received information in the reservation signal, each vehicle estimates which RBs are free and which ones are to be used by vehicles with pending packets with lower remaining life time $D(i,l)$

**Algorithm 5.3**: Distributed radio resource allocation

1. **Inputs:**

   $N_{RB}^{SL}$ : Number of RBs in SL

   $\alpha_j$ : Current mode of cluster j $\alpha_j \in \{0,1\}$

2. **If** $\alpha_j = 1$ { UEs within cluster j operate in sidelink Mode}
3. **for** each packet $l$ within queue of user $i$ do
4. Compute $D(i,l)$ for each packet generated by vehicle $i$ based on (5.8)
5. **End**
6. Call sensing RBs {Sense all resources within the selection window based on short-term sensing}.
7. Find RBs that are not reserved by another UE with lower life time.
8. **Set** $D_{TTI}^{x}(i,l)$ to 2
9. Compute $\beta_x(i,l)$ based on (5.10)
10. **If** Available RBs within $D_{TTI}^{x}(i,l) \geq \beta_x(i,l)$
11. Allocate RBs with best channel condition within $D_{TTI}^{x}(i,l)$ equal to 2 symbols vehicle $i$ with highest priority $(\mathrm{Min}\, D(i,l))$
12. **Else**
13. **Set** $D_{TTI}^{x}(i,l)$ to 4
14. **If** Available RBs within $D_{TTI}^{x}(i,l) \geq \beta_x(i,l)$
15. Allocate RBs with best channel condition within $D_{TTI}^{x}(i,l)$ equal to 4 symbols vehicle $i$ with highest priority $(\mathrm{Min}\, D(i,l))$ .
16. **Else**

    **Set** $D_{TTI}^{x}(i,l)$ to 7
17. Allocate RBs with best channel condition within $D_{TTI}^{x}(i,l)$ equal to 7 symbols vehicle $i$ with highest priority $(\mathrm{Min}\, D(i,l))$
18. **End**
19. **End**
20. Send the reservation signal to other UEs
21. Transmit packet $l$ of UE $i$
22. Remove packet $l$ from UE $i$ queue

23 **End**

computed as in (5.8). Then, the best RB in terms of the channel quality will be reserved by the vehicle, *i* to transmit the most urgent packet (i.e., the packet with lower life time) pending in the queue of vehicle *i*. For this purpose, the algorithm first needs to select the value of $D_{TTI}^{SL}(i,l)$ and obtain $\beta_{SL}(i,l)$, which are the RBs needed to transmit the packet *l* of vehicle *i* in SL Like in the centralized approach, the proposed scheme supports a flexible structure for dynamic scheduling of users with different TTI sizes using mini-slots of 2, 4 or 7 symbols in accordance to each user requirements. The algorithm starts with the $D_{TTI}^{SL}(i,l)$ equal to 2, and then computes the required RBs $\beta_{SL}(i,l)$ as follows.

$$\beta_{SL}(i,l) = \frac{S_m}{D_{TTI}^{SL}(i,l).\, SP_{eff,SL} \cdot B} \tag{5.10}$$

If $\beta_{SL}(i,l)$ is higher than the available resources, the algorithm will increase the TTI size until ensuring that all the resources requested by user *i* are met or until reaching the maximum TTI size of 7 symbols (lines 8-19).

### 5.4.3 Performance Evaluation for MS-RRAS

In this section, we evaluate the performance of the proposed mode selection and radio resource allocation solution through system level simulations performed in MATLAB.

#### 5.4.3.1 Simulation Setup

Our simulation model is based on a gNB that supports a cell with a channel organized in 80 RBs for UL and 80 RBs for DL composed by 12 subcarriers with subcarrier separation $\Delta f$=30 kHz, which corresponds to one of the 5G NR numerologies defined in [62]. The number of RBs in the UL is divided into $N_{RB}^{UL}$ =40 RBs for UL transmission (cellular mode) and $N_{RB}^{SL}$ =40 RBs for sidelink transmission. The model considers vehicular UEs communicating through cellular mode (uplink / downlink) and via sidelink (direct V2V). The users move along a 2-lane highway are assumed to enter the cell coverage following a Poisson process with arrival rate ($\lambda_v$). The highway is divided clusters by sectioning the road into smaller zones according to the length of the road. All relevant system and simulation parameters are summarized in Table 5.2.

The presented evaluation results intend to assess and illustrate the performance of the proposed a mode selection and resource allocation solution in terms of latency, packet success rate and RB utilization.

### 5.4.3.2 Benchmark Techniques

For performance comparison, we consider the following approaches to compare with our proposed mode selection and resource allocation algorithm (MS-RRAS):

- **Signal-Based mode selection with fixed TTI (SMS-FTTI):** This technique considers the MSBS strategy discussed in Subsection 4.4. In addition, for the resource allocation, we assume that the radio resources are allocated over a fixed time slot (i.e., 7 symbols).

- **Signal-Based mode selection with dynamic TTI (SMS-DTTI):** In order to see the impact of the dynamic allocation of radio resources, we consider the MSBS strategy discussed in Subsection 4.4 and resource allocation policy similar to SMS-FTTI but that supports dynamic allocation of resources over a fraction of a slot (i.e.,2, 4, 7 symbols) and takes into account only the quality of the links between V2V users in sidelink mode and between the base station and the vehicles in cellular mode.

- **SL-Fixed TTI (SL-FTTI):** This is a reference scheme in which the vehicles always operate in sidelink mode. As for the resource allocation, we assume that the radio resources are allocated over a fixed time slot (i.e., 7 symbols).

- **SL-Dynamic TTI (SL-DTTI):** This reference scheme assumes that the vehicles always operate in sidelink mode, but supports dynamic allocation of radio resources over mini-slots, i.e. considering 2, 4 or 7 symbols.

TABLE 5.2: Simulation parameters

<table>
<tr><th>Parameter</th><th colspan="2">Values</th></tr>
<tr><td>Cell radius</td><td colspan="2">1500m</td></tr>
<tr><td rowspan="3">Number of RBs per cell</td><td rowspan="2">$N_{UL}$ =80 RBs</td><td>$N_{RB}^{SL}$ =40 RBs</td></tr>
<tr><td>$N_{RB}^{SL}$ =40 RBs</td></tr>
<tr><td colspan="2">$N_{DL}$ =80 RBs</td></tr>
<tr><td>Frequency</td><td colspan="2">2.6 GHz</td></tr>
<tr><td>$\Delta f$</td><td colspan="2">30 kHz</td></tr>
<tr><td>Path loss model</td><td colspan="2">The path loss and the LOS probability for cellular mode are modeled as in [63]. In sidelink mode, all V2V links are modeled based on freeway case (WINNER+B1) [ITU-R] [64].</td></tr>
<tr><td>height of the gNB</td><td colspan="2">10m</td></tr>
<tr><td>Spectral efficiency model to map SINR.</td><td colspan="2">Model in section A.1 of [65]. The maximum spectral efficiency is 1 b/s/Hz.</td></tr>
<tr><td>Shadowing standard deviation</td><td colspan="2">3 dB in LOS and 4 dB in NLOS.</td></tr>
<tr><td>Base station antenna gain</td><td colspan="2">5 dB</td></tr>
<tr><td>Length of the highway</td><td colspan="2">3Km</td></tr>
<tr><td>Number of lanes</td><td colspan="2">2 in one direction</td></tr>
<tr><td>Lane width</td><td colspan="2">4 m</td></tr>
<tr><td>Size of cluster</td><td colspan="2">100, 200,300,400,600,1200m</td></tr>
<tr><td>Vehicular UE height</td><td colspan="2">1.5m</td></tr>
<tr><td>vehicle speed</td><td colspan="2">80 Km/h</td></tr>
<tr><td>Vehicle arrival rate $\lambda_v$</td><td colspan="2">Varied from 1 to 8 vehicles/s</td></tr>
<tr><td>Packet arrival rate $\lambda_a$</td><td colspan="2">1 packets/s</td></tr>
<tr><td>Message size ($S_m$)</td><td colspan="2">800 bytes</td></tr>
<tr><td>TTI duration</td><td colspan="2">2,4,7 symbols</td></tr>
<tr><td>Time window T</td><td colspan="2">3s</td></tr>
<tr><td>$\gamma_{TH}$</td><td colspan="2">12 dB</td></tr>
<tr><td>$W_{th}$</td><td colspan="2">30ms</td></tr>
</table>

### 5.4.3.3 Performance in terms of Latency

In this subsection, the performance of our proposed MS-RRAS in terms of the latency is compared with the reference benchmark schemes. Figure. 5.3 illustrates the average latency for V2V service. We can infer from Figure 5.3 that the average latency increases for all the approaches when vehicle arrival rate ($\lambda_v$) is increased because more vehicles will use the network and request more RBs to be used for the transmissions. This causes an increase in the waiting time and therefore increases the latency. From the presented results, we notice that the MS-RRAS in this chapter reduces the latency compared to the reference schemes. For the MS-RRAS, the average latency is around 18.8ms when the vehicle arrival rate is 8 vehicles/s, while for the SMS-FTTI approach, the latency is about 27.24ms (i.e. MS-RRAS achieves a relative gain of 30 %).

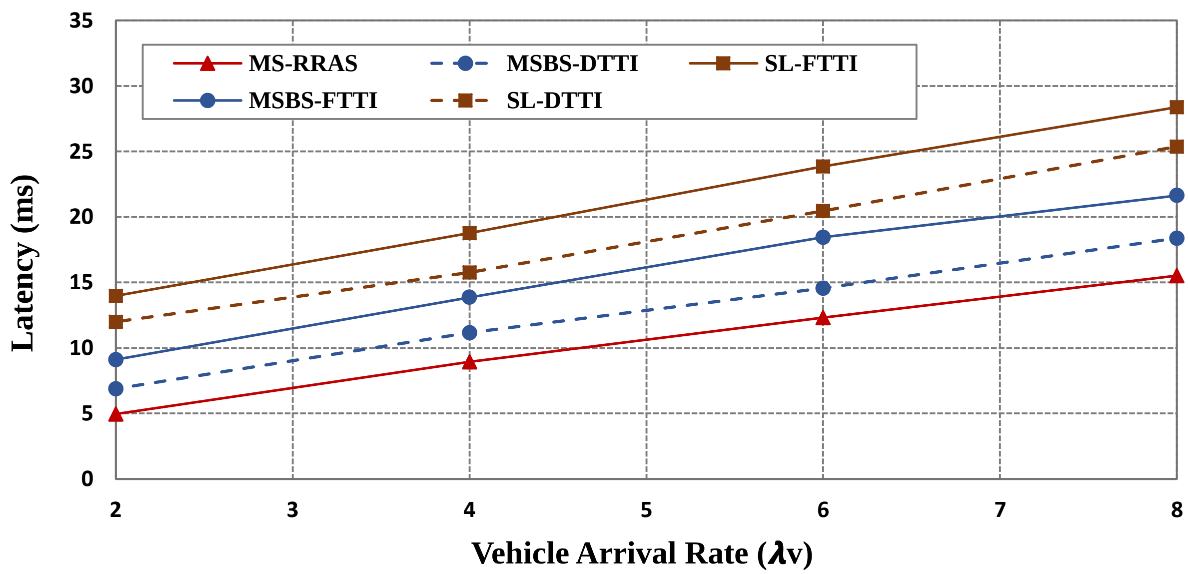


**Figure 5.3**: Average Latency as a function of the vehicle arrival rate $\lambda_v$ (vehicles/s).

Similarly, in case of the SMS-DTTI, the latency is only about 23.14ms when the vehicle arrival rate is 8 vehicles/s (i.e. MS-RRAS achieves a relative gain of 18%) and for the SL-FTTI approach, the latency is about 37.15ms (i.e. MS-RRAS achieves a relative gain of 49%). In case of the SL-DTTI, the latency is only about 33.12ms (i.e. MS-RRAS achieves a relative gain of 43 %). The gains are achieved because the proposed approach makes more efficient use of RBs. Thus it reduces the corresponding waiting time and the transmission delay.

Moreover, Figure 5.4 shows the effect of the cluster size on the latency by considering different cluster sizes at arrival rate of 6 vehicles/s. The figure illustrates the behavior of the MS-RRAS and the reference schemes. From Figure 5.4, we can see that the average latency increases for all the approaches when the cluster size increases (i.e., the distance between transmitter and receivers is increased). Here, we can observe that the MS-RRAS outperforms the reference schemes. For the MS-RRAS, when the cluster size is 200m, the average latency is around 8.7 ms. Instead, for the SMS-DTTI and SMS-FTTI reference approaches, the latency is 13.8ms and 18.76 ms, respectively. For the SL-DTTI and SL-FTTI strategies, the average latency is 23.8 ms and 27.2 ms, respectively.

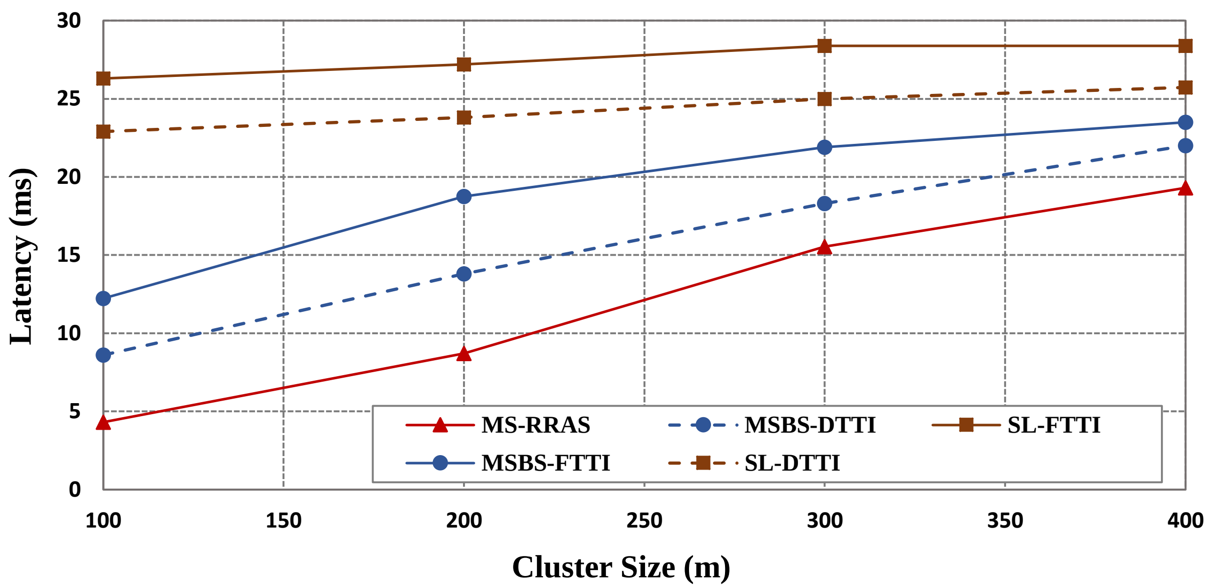


**Figure 5.4**: Average Latency as a function of the cluster size (m).

In Figure.5.5, we investigate the probability of exceeding the maximum delay (i.e., $W_{j,i}(l) > W_{TH}$) against the vehicle arrival rate ($\lambda_v$). As shown in the figure, increasing the traffic load leads to an increase in the probability of exceeding the maximum delay because more vehicles will use the network and request more RBs to be used for the transmissions. This causes an increase in the waiting time and therefore increases the latency. It can be also noted that the MS-RRAS can substantially reduce the probability of exceeding the maximum delay compared to the reference schemes.

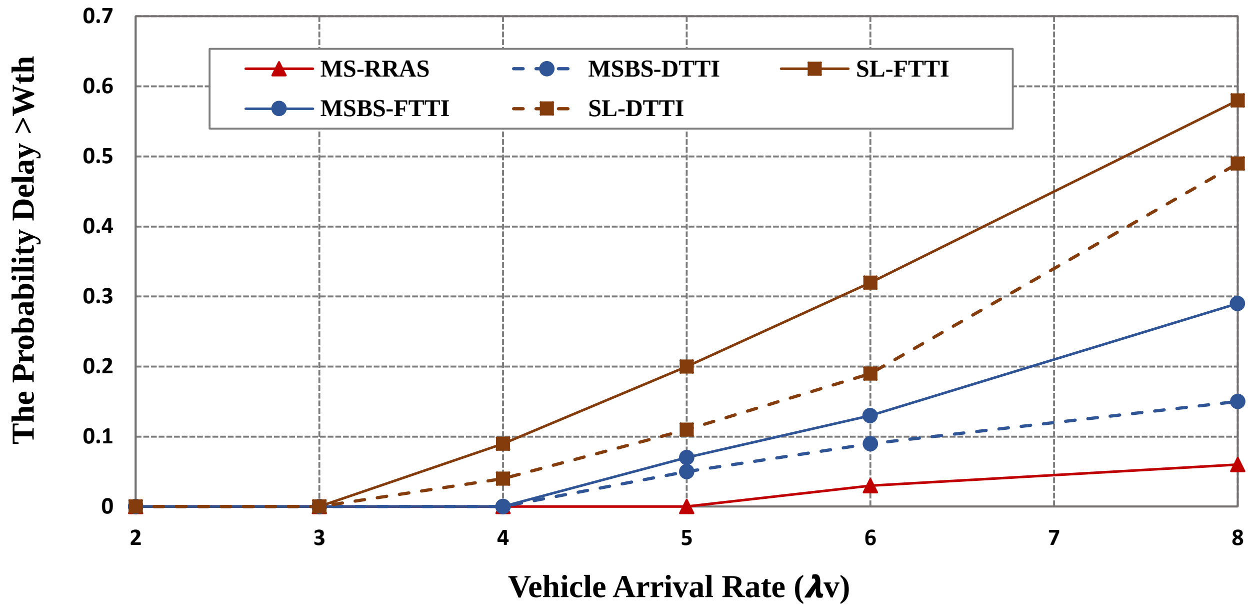


**Figure 5.5**: The probability of $W_{j,i}(l) > W_{TH}$ as a function of the vehicle arrival rate $\lambda_v$ (vehicles/s).

### 5.4.3.4 Performance in Terms of Packet Success Rate

In this subsection, the performance of the MS-RRAS is compared with the reference schemes in terms of packet success rate i.e. the number of packet delivered successfully normalized to the total number of packets that have been transmitted. Successful delivery means that packets are not erroneous and that they are delivered on time before the maximum threshold $W_{th}$. Figure 5.6 illustrates the behavior of the MS-RRAS and the reference schemes. Here, we can clearly observe that the MS-RRAS maintains a higher packet success rate than the reference schemes. Specifically, the MS-RRAS delivers around 99 % and 97 % of packets successfully when vehicle arrival rate is 5 vehicles/s and 6 vehicles/s, respectively. For the SMS-DTTI reference approaches, packet success rate is only about 95 % and 91 %, when the vehicle arrival rate is 5 vehicles/s and 6 vehicles/s, respectively. In case of the SMS-FTTI approach, the packet success rate ratio is only about 93 % and

87 %, respectively, when the vehicle arrival rate is 5 vehicles/s and 6 vehicles/s. For the SL-DTTI reference approach, packet success rate is only about 88% and 79 % when the vehicle arrival rate is 5 vehicles/s and 6 vehicles/s, respectively. While in case of the SL- FTTI reference approach, the packet success rate is only about is only about 80 % and 68 %, respectively, when the vehicle arrival rate is 5 vehicles/s and 6 vehicles/s.

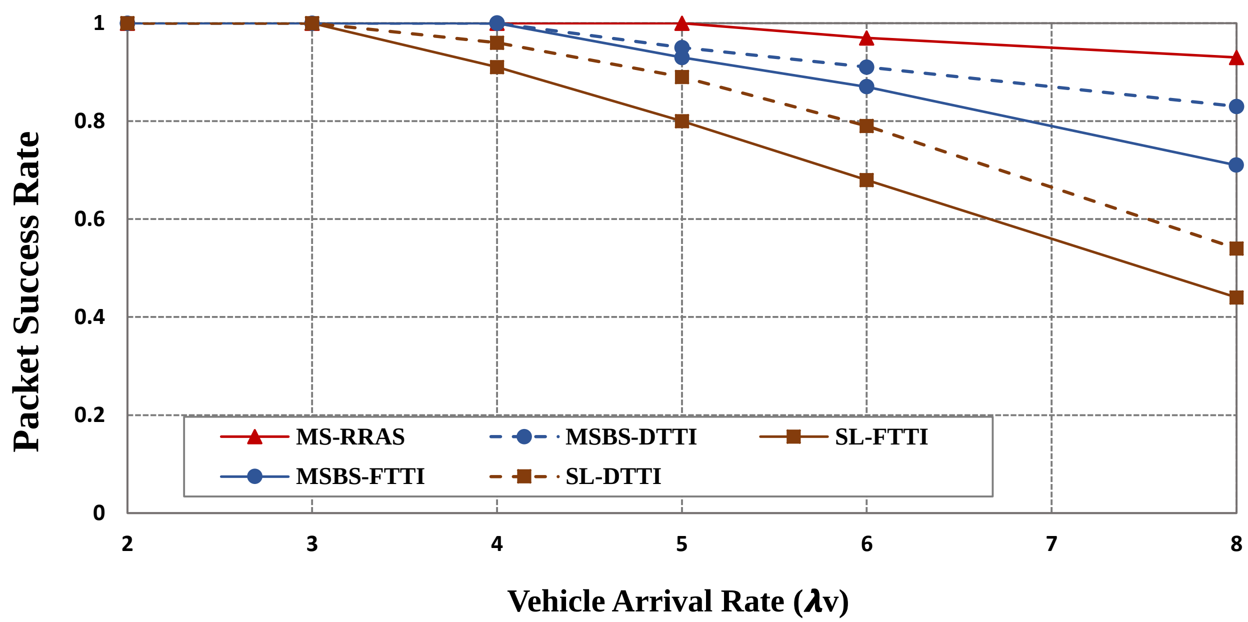


**Figure 5.6**: Packet success rate as a function of the vehicle arrival rate $\lambda_v$ (vehicles/s.

The gains achieved result from a more efficient utilization of radio resources. In addition, the proposed approach maintains a higher quality of signal to satisfy the SINR threshold constraint that improves the packet success rate.

Figure 5.7 compares the performance of the MS-RRAS in terms of packet success rate against the SMS-DTTI, the SL-FTTI and SL-DTTI schemes for different cluster sizes. From the figure, we observe that the packet success rate decreases with the cluster size, because the channel gain of V2V links decreases when the cluster size increases. Again, the proposed scheme outperforms the reference schemes.

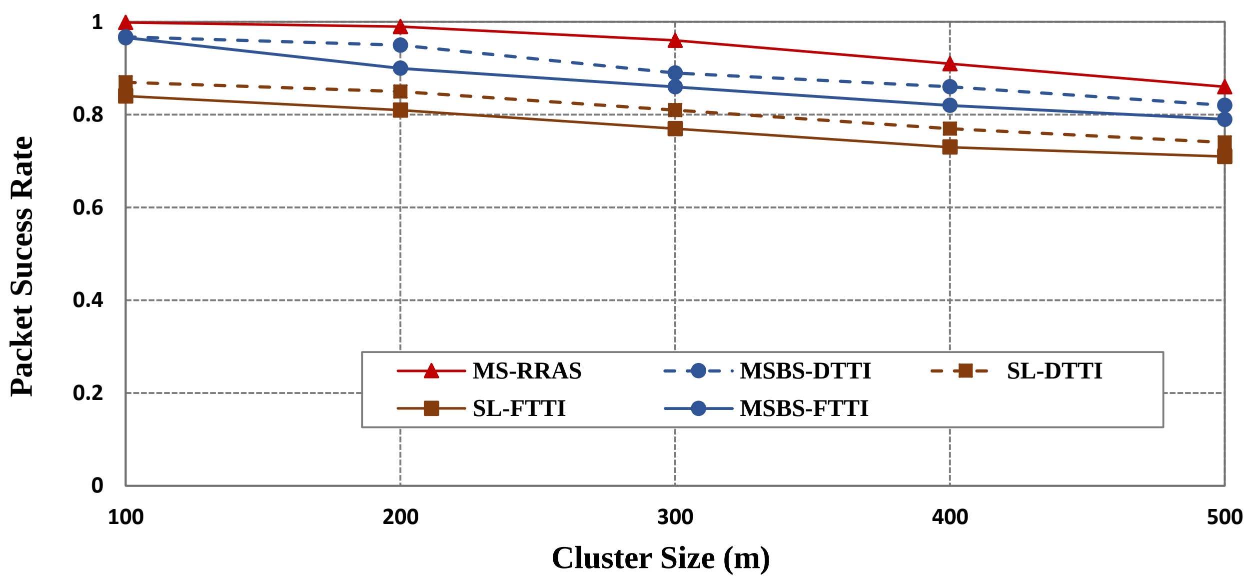


**Figure 5.7**: Packet success rate as a function of cluster size(m).

### 5.4.3.5 Performance in Terms of RBs Utilization

In this subsection, the performance of the MS-RRAS is compared with the reference schemes in terms of the obtained RB utilization, i.e. the number of used RBs normalized to the number of total available RBs, in UL and DL, respectively. Figure 5.8 and Figure 5.9 present the obtained RB utilization for UL and DL, as a function of the vehicle arrival rate ($\lambda_v$). Since SL and UL make use of the same set of RBs, the results included in Fig.5.8 refer to the total utilization by both links for V2V.

From the presented results, we notice that the proposed MS-RRAS maintains high resource utilization compared to the reference schemes in different load scenarios. For the MS-RRAS, when the vehicle arrival rate is 4 vehicles/s, the system utilizes around 88 % and 44 % of radio resources in uplink and downlink, respectively. For the SMS-DTTI reference approach, the utilization is about 69 % in uplink (i.e. MS-RRAS achieves a relative gain of 27%) and 34 % in downlink (i.e. MS-

RRAS achieves a relative gain of 29%). Similarly, in case of the SMS-FTTI, the utilization is only about 63 % in uplink (i.e. MS-RRAS achieves a relative gain of 39%) and 31 % in downlink (i.e. MS-RRAS achieves a relative gain of 41%). This is because the proposed strategy inherently tackles the network traffic load dynamics in each cluster of each cluster before making a selection decision (i.e., the MS-RRAS distributes the load on a balanced basis between the two modes). This leads to the effective use of resources and avoid lack of resources. In case of the SL-DTTI strategy, the utilization is only about 50 % of radio resources in uplink (i.e. MS-RRAS achieves a relative gain of 76 %) and 25 % in downlink (i.e. MS-RRAS achieves a relative gain of 76%) when the vehicle arrival rate is 4 vehicles/s. While in case of the SL-FTTI strategy, the utilization is only about 42 % of radio resources in uplink (i.e. MS-RRAS achieves a relative gain of 109 %) and 21% in downlink (i.e. MS-RRAS achieves a relative gain of 109%) when the vehicle arrival rate is 4 vehicles/s. This is because in the SL-FTTI and SL-DTTI strategies all cluster will operate in SL mode and due to the limited available resources (i.e., 40 RBs), the SL will experience lack of resources.

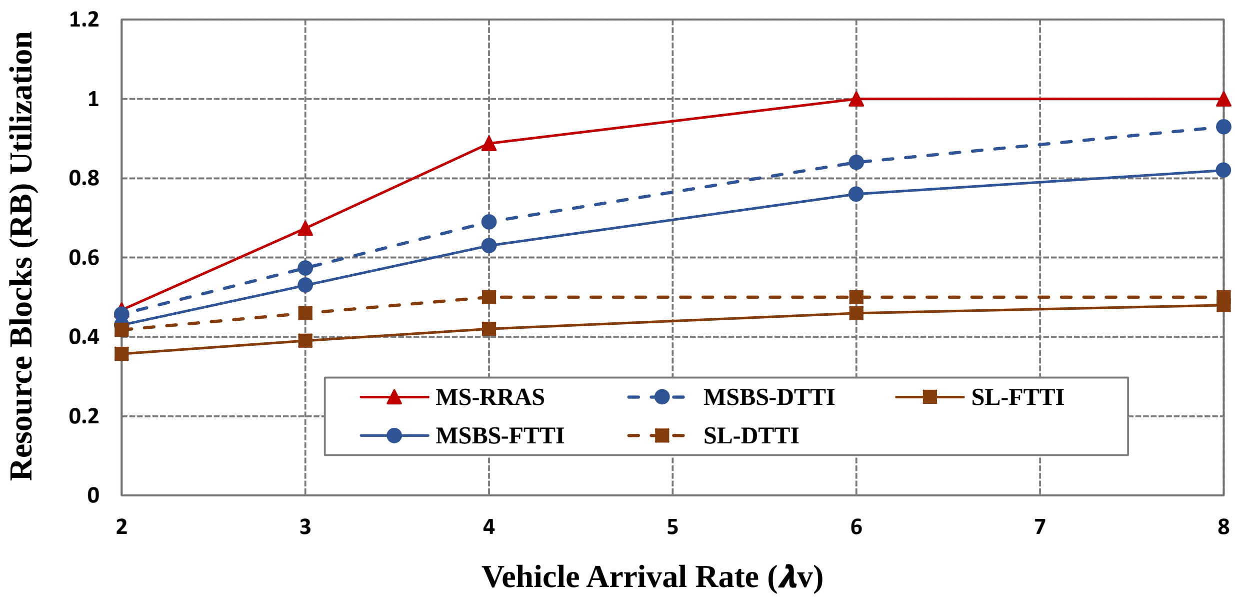


**Figure 5.8**: Uplink RB utilization as a function of the vehicle arrival rate $\lambda_v$ (vehicles/s).

Figure 5.10 compares the performance of the MS-RRAS against the reference schemes in uplink for different cluster sizes at arrival rate of 6 vehicles/s. We notice that the obtained RB utilization

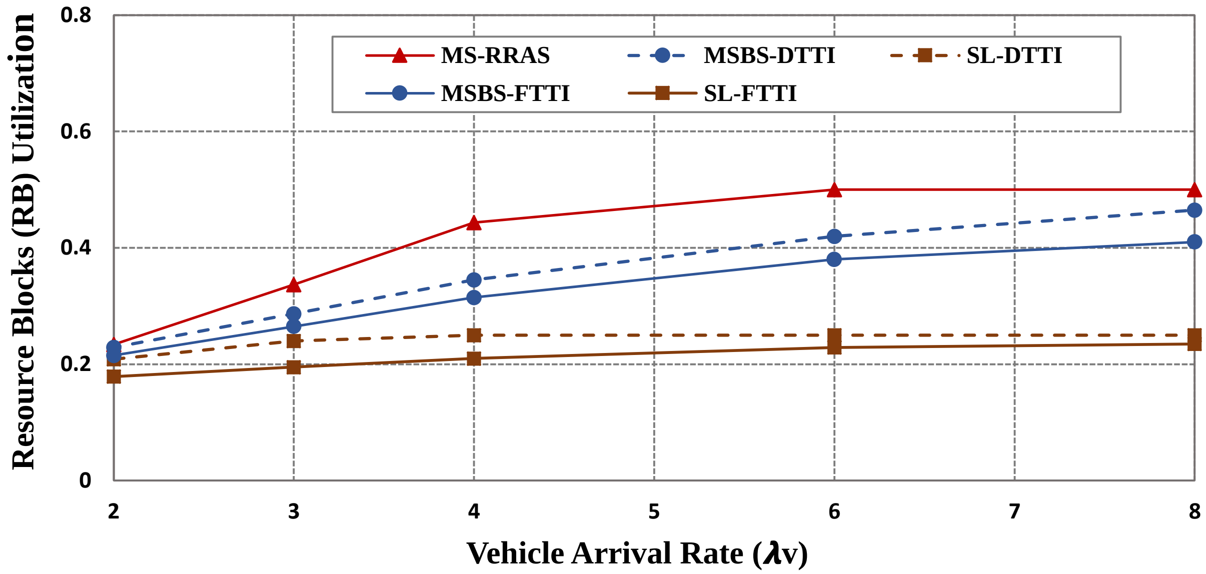


**Figure 5.9**: Downlink RB utilization as a function of the vehicle arrival rate $\lambda_v$ (vehicles/s).

decreases for all the approaches when cluster size is larger (i.e., the distance between transmitter and receivers is increased). This is because in the case of the MS-RRAS, as distance increases some clusters will remain in the same mode in which they operate because the minimum service requirements, i.e., acceptable delay and available radio resources, are not met in the other mode. While in SMS-DTTI and SMS-FTTI, as distance increases, more clusters will select cellular mode as a preferred mode of operation because a better signal will be received. This will lead to lack of resources as the total amount of resources for each mode should not exceed the total amount of available resources i.e., 40 RBs. Here, we can also observe that the proposed strategy outperforms the references schemes because the proposed approach not only takes into account the signal received, but also ensures that the required RBs by each cluster is less than the number of RBs available for the mode in which the vehicles are to be switched and operated. Specifically, in the MS-RRAS, when the size of cluster is 300m, the system utilizes around 82 % of radio resources in uplink. In turn, in case of the SMS-FTTI and SMS-DTTI strategies, the utilization of radio resources in uplink is only about 60 % and 66 %, respectively (i.e. proposed approach achieves a relative gain of 0.36 and 24%, respectively). For the SL-FTTI and SL-DTTI reference approaches, the utilization is only about 46% and 50 % in uplink, respectively (i.e. proposed approach achieves a relative gain of 78% and 64%, respectively).

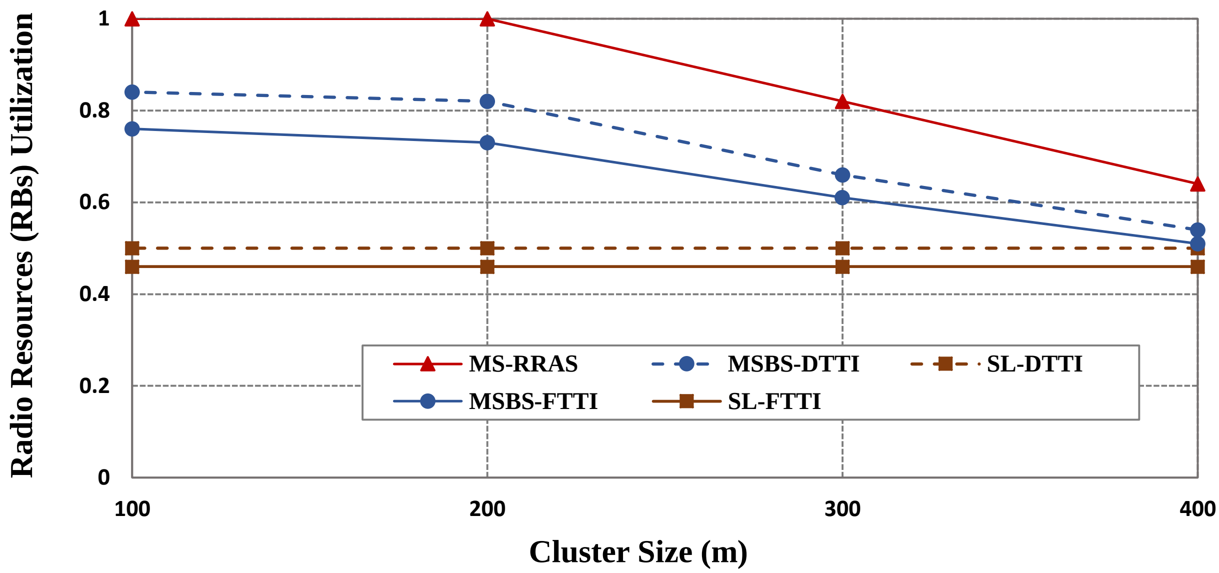


**Figure 5.10**: Uplink RB utilization as a function of cluster size (m)

## 5.5 Concluding Remarks

In this chapter, we have proposed a novel mode selection strategy for V2V communications with the target of minimizing the probability of exceeding the maximum delay under the constraint of satisfying the reliability requirement. The proposed approach takes into account three different aspects: the quality of the links between V2V users in sidelink mode and between the base station and the vehicles in cellular mode, the available resources, and the latency. In addition, we have proposed two resource allocation strategies, one centralized for cellular mode and the other distributed for sidelink mode. The proposed approaches support dynamic allocation of radio resources over mini-slots.

Extensive simulations were conducted to validate and analyse the performance of our proposed solution. This strategy has been compared against reference approaches that consider only the quality of the links and support allocation of radio resources over fixed time slots (i.e., SMS-FTTI) and over dynamic mini-slot durations (i.e., SMS-DTTI). In addition, the proposed strategy has been compared against other reference approaches in which the vehicles always operate in sidelink mode, considering also fixed (i.e. SL-FTTI) or dynamic (i.e. SL-DTTI) mini-slot durations.

Simulation results have shown the capability of the proposed algorithms to allocate the resources efficiently and improve the network performance in terms of delay, packet success rate, and resource utilization. From the presented results, we notice that our proposed approach outperforms the reference schemes in terms of the latency (i.e. the proposed approach achieves a relative gain of up to 18 % and 30 % compared to the SMS-DTTI and SMS-FTTI reference approaches, respectively, and up to 43 % and 49 % compared to the SL-DTTI and SL-FTTI reference approaches, respectively). Moreover, results have also demonstrated an improvement in the packet success rate (e.g., packet success rate increases from 87%, 91%, 68% and 79% for SMS-FTTI, SMS-DTTI, SL-FTTI and SL-DTTI reference approaches, respectively, to 97 % for the proposed approach, when the vehicle arrival rate is 6 vehicles/s).

# Chapter 6

# Conclusions and future work

In this chapter, the main contribution of the thesis is summarized and the derived conclusions are discussed in Section 6.1. Additionally, directions for future research on the topics studied in this thesis are presented in Section 6.2.

## 6.1 Conclusions

The integration of V2V communication into cellular networks is a promising solution for enhancing the performance of V2X in cellular systems. By allowing direct communication between the mobile users, we can offload base stations and improve performance metrics such as network coverage, end-to-end latency, energy consumption, and spectral efficiency. In this thesis, we addressed some of the challenges arising from the integration of V2V communication in cellular systems and validated the potential of this technology by providing appropriate resource management solutions. Our main contributions have been in the context of radio access network slicing, mode selection, and radio resource allocation mechanisms.

With regard to the first research direction that focuses on the RAN slicing management, presented in Chapter 3, a novel strategy based on offline Q-learning and softmax decision-making has been proposed as an enhanced solution to determine the adequate split of resources between a slice for eMBB communications and a slice for V2X. The adopted Q-Learning algorithm tracks all possible actions' (i.e. resource splits) the system can take through an exploration-exploitation process in order to select the appropriate one. Then, starting from the outcome of the off-line Q-learning algorithm, a low-complexity heuristic strategy has been proposed to allow balancing the trade-off between achieving a fine granularity when setting the slicing ratios and keeping a moderate number of actions in the RL algorithm that facilitates the convergence of the algorithm in a reduced time, and achieving further improvements in the use of resources. The proposed solution has been compared against proportional and fixed reference schemes. The extensive performance assessment have revealed the ability of the proposed algorithms to improve network performance compared to the reference schemes, especially in terms of resource utilization, throughput, latency and outage probability.

Regarding the second research direction that focuses on the mode selection, two different mode selection solutions referred to as MSSB and MS-RBRS strategies have been proposed in chapter 4 for V2V communication over a cellular network. The MSSB strategy decides when it is appropriate to use one or the other mode, i.e. sidelink or cellular, for the involved vehicles, taking into account the quality of the links between V2V users, the available resources, and the network traffic load situation. Moreover, the MS-RBRS strategy not only selects the appropriate mode of operation but also decides efficiently the amount of resources needed by V2V links in each mode and allows reusing RBs between different SL users while guaranteeing the minimum interference requirements. A simulation-based analysis has been presented to assess the performance of the proposed strategies. The conducted simulations have revealed that the MS-RBRS and MSSB strategies are beneficial in terms of throughput, radio resource utilization, outage probability and latency under different offered loads comparing to the reference scheme. Specifically, from the results presented in chapter 4, it has been obtaiend that further improvements can be obtained by the MS-RBRS strategy. This is because the proposed solution ensures more RBs which can be used to serve the V2V traffic by taking advantage from the reuse of RBs for other SL V2V links that meet the interference requirements.

Last, we have focused on the resource allocation problem including jointly mode selection and radio resource scheduling. For the mode selection, a novel mode selection has been presented to decide when it is appropriate to select sidelink mode and use a distributed approach for radio resource allocation or cellular mode and use a centralized radio resource allocation. It takes into account three aspects: the quality of the links between V2V users, the available resources, and the latency. As for the radio resource allocation, the proposed approach includes a distributed radio resource allocation for sidelink mode and a centralized radio resource allocation for cellular mode. The proposed strategy supports dynamic assignments by allowing transmission over mini-slots. A simulation-based analysis has shown that the proposed strategies improved the network performance in terms of latency of V2V services, packet success rate and resource utilization under different network loads. The gains are achieved because the proposed solution makes more efficient use of RBs. Thus it reduces the corresponding waiting time and the transmission delay. In addition, the proposed solution maintains a higher quality of signal to satisfy the SINR threshold constraint that improves the packet success rate.

## 6.2 Directions for future work

The technical aspects of our contributions conducted in the context of this research open new horizons for future works. Some of the possible interesting directions for future research on the issues that our work has not yet covered are discussed as follows.

(i) Regarding the network slicing management, our study has focused on a RAN slicing framework based on an offline reinforcement learning followed by a low-complexity heuristic algorithm, which allocates radio resources to different slices (i.e., eMBB and V2X). An important extension is to consider more resource splits in order to optimally determine the most adequate one that improves the overall use of resources under the constraints of meeting the resource requirements for the users of different slices. In addition, Moreover, we focused only on the RAN slicing problem. It would be interesting to design practical network slicing strategies comprising both RAN and CN in the multi-cell with different dynamic traffic scenario.

(ii) Regarding the resource management including mode selection and radio resource allocation in V2V communication, we have focused on the V2V data exchange over cellular network. However, 5G will deploy existing and novel 3GPP (4G LTE, 5G NR) and non-3GPP (e.g., IEEE 802.11) RATs. In the 5G V2X context, the usage of multiple RATs may boost V2I/V2N network throughput and capacity. Therefore, one interesting extension would be to develop mechanisms and procedures that allow optimized integration between different RATs in order to extend the cellular network coverage and capacity, and reduce latency, considering heterogeneous scheduling and traffic flow management with various traffic flows for V2X communication.

(iii) From the prospective of mobility management, given the heterogeneity and complexity of the 5G scenarios, it is vital to study the appropriate mobility management for different V2X scenarios. Indeed, there is a need to propose novel flexible and adaptive handover mechanisms (i.e., handovers between different access networks) for V2X communication to guarantee seamless handover in single and multi-RATs and fulfill the requirements for low latency and high reliability.

Concluding, this thesis has addressed some important issues of cellular V2X communications in relation to radio access network slicing, mode selection, and radio resource allocation. It should be noted that the concerns mentioned so far are indicative challenges that the new technologies pose in the upcoming V2X capabilities in cellular networks. The network slicing, mode selection, and radio resource allocation management strategies we have proposed in this thesis cannot claim to be absolute and unique solutions to the problems under study. Nevertheless, we believe that the presented strategies can be a valuable contribution to the improvement of V2X over cellular networks and that our study can provide insights towards the design of efficient resource management techniques that leverage the capabilities of the future cellular V2X networks and are able to ensure the provision of high-quality services to the end users.

# References


[1] 5G Americas. “V2X Cellular Solutions” [Online]. October 2016. . http://www.5gamericas.org/files/2914/7769/1296/5GA_V2X_Report_FINAL_for_upload.pdf.

[2] A. Vinel, “3GPP LTE versus IEEE 802.11p/WAVE: Which technology is able to support cooperative vehicular safety applications?” IEEE Wireless Communications Letters, vol. 1, no. 2, pp. 125–128, Feb. 2012.

[3] S. Kato, M. Hiltunen, K. Joshi, and R. Schlichting, “Enabling vehicular safety applications over LTE networks,” in Proc. International Conference on Connected Vehicles and Expo (ICCVE), Las Vegas, NV, USA, Dec. 2013.

[4] Report ITU-R M.2412-0, “Minimum requirements related to technical performance for IMT-2020 radio interface(s)”, ITU-R, 2017.

[5] Study on New Radio (NR) Access Technology (Release 15), document 3GPP TR 38.912 V15.0.0, Jun. 2018.

[6] NGMN Alliance. Description of Network Slicing Concept. Accessed: Apr. 5, 2019. [Online]. Available: https://www. ngmn.org/_leadmin/user_upload/160113_Network_Slicing_v1_0.pdf

[7] 5G Americas, “Network Slicing for 5G Networks and Services,” Tech.Rep.,November,2016.[Online].Available: http://www.5gamericas.org/files/3214/7975/0104/5GAmericas Network Slicing 11.21 Final.pdf

[8] 3GPP TS 22.261 v16.0.0, “Service requirements for the 5G system; Stage 1 (Release 15)”, June 2018.

[9] 3GPP TS 28.530 v0.3.0 “Management of network slicing in mobile networks; Concepts, use cases and requirements (Release 15)”, August, 2018.

[10] 3GPP TS 23.501 v1.0.0, “System Architecture for the 5G System; Stage 2 (Release 15)”, Sep. 2018.

[11] Ericsson, “White Paper: 5G radio access - capabilities and technologies,”.2016.[Online].Available: https://www.ericsson.com/en/white-papers/5g-radio-access--apabilities-and-technologies.

[12] 3GPP TR 22.885, ” Study on LTE support for Vehicle to Everything (V2X) services ”, v14.0.0, Dec. 2015.

[13] ETSI TC ITS, Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Definitions, Std. ETSI TR 102 638 V1.1.2, 2015.

[14] 3GPP TR 22.886, ”Study on enhancement of 3GPP Support for 5G V2X Services”, v15.1.0, Mar. 2017.

[15] 3GPP TR 38.885 ,”*Study on NR Vehicle-to-Everything ”, (Release 16)*, document v16.0.0, Mar. 2019.

[16] Z. Li, M. Moisio, M. Uusitalo, P. Lunden, C. Wijting, F. Sanchez Moya, A. Yaver, and V. Venkatasubramanian, “Overview on initial METIS D2D concept,” in *1st International Conference on 5G for Ubiquitous Connectivity*, pp. 203–208, November 2014.

[17] D. Câmara and N. Nikaein, *Wireless Public Safety Networks Volume 1:Overview and Challenges*. Elsevier Science, 2015.

[18] K. S. Ali, H. ElSawy, and M. S. Alouini, “Modeling Cellular Networks With Full-Duplex D2D Communication: A Stochastic Geometry Approach,” *IEEE Transactions on Communications*, vol. 64, pp. 4409–4424, October 2016.

[19] S. Andreev, A. Pyattaev, K. Johnsson, O. Galinina, and Y. Koucheryavy, “Cellular traffic offloading onto network-assisted device-to-device connections,” IEEE Communications Magazine, vol. 52, pp. 20–31, April 2014.

[20] B. Soret, P. Mogensen, K. I. Pedersen and M. C. Aguayo-Torres,“Fundamental tradeoffs among reliability, latency and throughput incellular networks”, IEEE Globecom, Dec. 2014.

[21] 3GPP TS. 36.321, “Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC) protocol specification (v14.3.0, Release 14),” June 2017.

[22] Bazzi A., et al, ‘On the performance of IEEE 802.11 p and LTE-V2V for the cooperative awareness of connected vehicles’, IEEE Trans. Vehicular Tech., 66, (11), Nov. 2017.

[23] Di B., et al, ‘Non-orthogonal multiple access for high-reliable and lowlatency V2X communications in 5G systems’, IEEE JSAC, July 2017.

[24] Zhang X., et al, ‘Research on overlay D2D resource scheduling algorithms for V2V broadcast service’, IEEE VTC, Sept. 2016.

[25] Cecchini,G.,‘LTEV2Vsim:An LTE-V2V simulator for the investigation of resource allocation for cooperative awareness’, IEEE Conf. Models and Technologies for ITS, June 2017.

[26] 3GPP TR 23.287,” Architecture enhancements for 5G System (5GS) to support Vehicle-to-Everything (V2X) services ”, (Release 16), document v2.0.0, Sep. 2019.

[27] *S. Chen et al., “Technical Innovations Promoting Standard Evolution: From TD-SCDMA to TD-LTE and Beyond,” IEEE Wireless Commun., vol. 19, no. 1, Feb. 2013, pp. 60–66.*

[28] *ETSI GS MEC-IEG 005 V1.1.1, “Mobile-Edge Computing (MEC); Proof of Concept Framework,” Aug. 2015; http://www.etsi.org/deliver/ etsi_gs percent5CMEC-IEG percent5C001_099 percent5C005 percent5C01.01.01_60 percent5Cgs_MECIEG005v010101p.pdf).*

[29] *5G automotive vision,” Tech. Rep., Oct. 2015. [Online]. Available: 5g-ppp.eu/wp-ontent/uploads/2017/03/ 5GPPP-White-Paper-on-Automotive-Vertical-Sectors.pdf.*

[30] *3GPP TR 22.186 V16.2.0: Service requirements for enhanced V2X scenarios (Release 16), 3GPP Std., june 2019.*

[31] ETSI TC ITS, Intelligent Transport Systems; Vehicular Communications; Basic Set of Applications; Part 2: Specification of Cooperative Awareness Basic Service, Std. ETSI EN Std 302 637-2 V.1.3.0, 2013.

[32] ETSI TC ITS, Intelligent Transport Systems; Vehicular Communications; Basic Set of Applications; Part 3: Specification of Decentralized Environmental Notification Basic Service, Std. ETSI EN Std 302 637-3 V.1.2.0, 2013.

[33] *3GPP TS 23.285 version 14.2.0 Release 14, “Architecture enhancements for V2X services ,” Tech. Rep., May. 2017.*

[34] 3GPP TS 36.213, “Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures (v15.3.0, Release 15),”, Jun. 2019.

[35] GSMA - Network Slicing Use Case Requirements (54pgs). Accessed: Aug. 18, 2018. [Online]. Available: https://www.gsma.com/futurenetworks/5g/network-slicing-use-caserequirements- whitepaper.

[36] Minimum Technical Performance Requirements for IMT-2020 radio inter- face(s), document ITU-R M.2410-0, Nov. 2017.

[37] 3GPP TR 28.801 V1.2.0, “Study on management and orchestration of network slicing for next generation network (Release 15)”, May 2017.

[38] 5G PPP Architecture Working Group. (Dec. 2017). View on 5G Architecture. [Online]. Available: https://5g-ppp.eu/wp-content/uploads/ 2017/02/5G-PPP-5G-ArchitectureWP-For-public-consultation.pdf.

[39] METIS-II/D2.4 White Paper:” Final Overall 5G RAN Design,"’, June, 2017. Accessed: sep. 24, 2018. [Online]. Available: https://metis-ii.5g-ppp.eu/wp-content/uploads/deliverables/METIS-II_D2.4_V1.0.pdf

[40] 5GCARS. “Final 5G V2X Radio Design” [Online]. May 2019. https://5gcar.eu/wp-content/uploads/2019/06/5GCAR_D3.3_v1.0.pdf.

[41] Ericsson, “Network functions virtualization and software management”, Dec. 2014.

[42] GSMA., An Introduction to Network Slicing. Accessed: Feb. 6, 2019. [Online]. Available: https://www.gsma.com/futurenetworks/wp-content/uploads/2017/11/GSMA-An-Introduction-to-Network-Slicing.pdf.

[43] Ericsson, “White Paper: 5G radio access - capabilities and technologies,”.2016.[Online].Available: https://www.ericsson.com/en/white-papers/5g-radio-access--apabilities-and-technologies.

[44] Campolo C, Molinaro A, Iera A, Menichella F (2017) “5G network slicing for vehicle-to-everything services”, IEEE Wireless Communications,24(6):38–45.

[45] 3GPPTR23.786 v0.8.0. Study on architecture enhancements for EPS and 5G system to support advanced V2X services, Rel. 16,Aug.2018.

[46] H. Zhang, Na Liu, X. Chu, K. Long, A. Aghvami, Victor C. M. Leung: Network Slicing Based 5G and Future Mobile Networks: Mobility, Resource Management, and Challenges. IEEE Communications Magazine 55(8): 138-145 (2017).

[47] M. R. Sama, X. An, Q. Wei, and S. Beker, ``Reshaping the Mobile core network via function decomposition and network slicing for the 5G era,"in Proc. IEEE Wireless Commun. Netw. Conf. Workshops (WCNCW), Apr. 2016, pp. 90_96.

[48] X. Foukas, N. Nikaein, M. Kassem, M. K. Marina, and K. Kontovasilis “FlexRAN: A Flexible and Programmable Platform for Software-Defined Radio Access Networks”, International Conference on emerging Networking EXperiments and Technologies, Irvine, California, USA. Dec. 2016.

[49] V. Sciancalepore, K. Samdanis, X. Costa-Perez, D. Bega, M. Gramaglia, A. Banchs, “Mobile Traffic Forecasting for Maximizing 5G Network Slicing Resource Utilization”, IEEE INFOCOM 2017.

[50] H. M. Soliman, A. Leon-Garcia, “QoS-Aware Frequency-Space Network Slicing and Admission Control for Virtual Wireless Networks”, IEEE GLOBECOM, 2016.

[51] D. Thai Hoang1, D. Niyato1,and P. Wang “Optimal Cross Slice Orchestration for 5G Mobile Services,” arXiv:1712.05912v1, 16, Dec. 2017.

[52] A. Aijaz, “Hap-SliceR: A Radio Resource Slicing Framework for 5G Networks with Haptic Communications”, IEEE Systems Journal, 2017.

[53] M. Jiang, M. Condoluci, T. Mahmoodi, “Network slicing in 5G: an Auction-based model”, IEEE ICC 2017.

[54] P. Caballero, A. Banchs, G. de Veciana, X. Costa-Pérez, “Network Slicing Games: Enabling Customization in Multi-Tenant Networks”, IEEE INFOCOM, 2017.

[55] M. I. Kamel, Long Bao Le, Andr´e Girard, , “LTE Wireless Network Virtualization: Dynamic Slicing via Flexible Scheduling”, IEEE 2014.

[56] Eurecom, “Openairinterface 5G Wireless implementation, ” 2017. [Online]. Available: https://gitlab.eurecom.fr/ oai/open airinterface5g.

[57] F. Bahlke, O. D. Ramos-Cantor, S. Henneberger , and M.. Pesavento, ‘‘Optimized Cell Planning for Network Slicing in Heterogeneous Wireless Communication Networks,’’ IEEE Communications Letters 22(8): 1676-1679 (2018).

[58] L. Tang, QI Tan , Y. Shi, C. Wang , and A. Qianbin Chen, ‘‘Adaptive Virtual Resource Allocation in 5G Network Slicing Using Constrained Markov Decision Process,’’ IEEE Access, vol. 6, , pp. 31184–61195, Oct. 2018.

[59] V. Sciancalepore, L. Zanzi, Xavier Costa-Perez, and Antonio Capone, “ONETS: Online Network Slice Broker From Theory to Practice”,10 Jan, 2018 . [online]. Available: https://arxiv.org/ pdf/1801.03484.pdf

[60] R. S. Sutton and A. G. Barto, Reinforcement Learning: An Introduction. Cambridge, MA, USA: MIT Press, 1998.

[61] X. Zhou, R. Zhang, and C. K. Ho, "Wireless information and power transfer: Architecture design and rate-energy tradeoff", IEEE Transactions on Communications, vol. 61, no. 11, pp. 4754–4767, November 2013.

[62] 3GPP TS 38.211 v15.2.0 "NR; Physical channels and modulation (Release 15)", June, 2018.

[63] Report ITU-R M.2135 "Guidelines for evaluation of radio interface technologies for IMT-Advanced", 2009.

[64] WINNER II Channel Models, D1.1.2 V1.2., available at http://www.cept.org/files/1050/documents/winner2%20%20final%20r report.pdf.

[65] 3GPP TR 36.942 v15.0.0, "Radio Frequency (RF) system scenarios", September, 2018.

[66] G. Yu, L. Xu, D. Feng, R. Yin, `` Joint Mode Selection and Resource Allocation for Device-to-Device Communications," IEEE Trans. Commun., vol. 62, no. 11, pp. 3814_3824, Nov. 2014.

[67] H. Chou and R. Y. Chang, `` Joint Mode Selection and Interference Management in Device-to-Device Communications Underlaid MIMO Cellular Networks," IEEE Trans. Commun., vol. 16, no. 2, pp. 1120_1134, Feb. 2017.

[68] Y. Dai, M. Sheng, and X. Shen, `` Joint Mode Selection and Resource Allocation for D2D-Enabled NOMA Cellular Networks ," IEEE Trans. On Vehicular., vol. 68, no. 7, pp. 6721_ 6733, Jul. 2019.

[69] R. Wang, Jun Zhang, S.H. Song, and K. B. Letaief, "QoSAware Joint Mode Selection and Channel Assignment for D2D Communications," in IEEE ICC 2016 - Wireless Communications Symposium.

[70] R. Ma, N..Xia, H. Chen, C. Chiu, and Chu-Sing Yang , "Mode Selection, Radio Resource Allocation, and Power Coordination in D2D Communications ," IEEE Wireless Communications, Jun. 2017.

[71] Yi Li, M. Cenk Gursoy, Senem Velipasalar, Jian Tang, “Joint Mode Selection and Resource Allocation for D2D Communications via Vertex Coloring,” arXiv:1708.00872v1. 2 Aug 2017

[72] Y. Yang, Qingyi Quan, Junfei Li, Tao Peng, Wenbo Wang “Network Assisted Mode Selection in Multi-Cell for D2D Communication Underlaying LTE-A System,” in IEEE 2013.

[73] A. Omri, and M. O. Hasna, `` A Distance-Based Mode Selection Scheme for D2D-Enabled Networks with Mobility ,'' IEEE Trans. Commun., vol. 17, no. 7, pp. 4326_4340, Jul. 2018.

[74] G. Wang, S. Zhou, Z. Niu , `` Mode Selection in UAV-aided Vehicular Network: an Evolutionary Game Approach ,'' in Proc. IEEE Wireless Commun. Conf., 2018.

[75] P. Zeng, X. Li , `` A Strategy of Reusing Mode Selection for Device to Device Communication,'' 14th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), 2018. pp. 1231-1236.

[76] Ch. Chen , , Chi. Sung, and H.Chen , `` Capacity Maximization Based on Optimal Mode Selection in Multi-Mode and Multi-Pair D2D Communications Networks with Mobility ,'' IEEE Trans. On Vehicular., vol. 68, no. 7, pp. 5624_5634, Jul. 2019.

[77] A. Ortiz, A. Asadi, and M. Engelhardt, `` CBMoS: Combinatorial Bandit Learning for Mode Selection and Resource Allocation in D2D Systems,'' IEEE jour. On selected areas in communications., vol. 37, no. 10, pp. 2225_ 2238, Oct. 2019.

[78] Lei Qingyun Hao,Zhangdui Zhong, “Mode Selection and Resource Allocation in Device-to-Device Communications With User Arrivals and Departures”. IEEE Access : 5209-5222 (2016).

[79] Wanlu Sun ,On Medium Access Control for Vehicular Communication over Device-to Device Links: Radio Resource Management and Network Synchronization,2016

[80] 3GPP, “Study on Single-Cell Point-To-Multipoint transmission for EUTRA,” TR 36.890, 3rd Generation Partnership Project (3GPP), June 2015.

[81] R1-133140, “Cluster-based unified approach to D2D direct communication”, ITRI RAN1_74, Chicago, USA, April 2013.

[82] M. Gharbieh, A. Bader, and H. ElSawy, “Self-Organized Scheduling Request for Uplink 5G Networks: A D2D Clustering Approach,” IEEE Trans.Veh. Technol., vol. 67, no. 2, pp 1197-1209, Oct. 2018.

[83] X. Lin, J. G. Andrews, and A. Ghosh, "Spectrum sharing for device- to- device communication in cellular networks," IEEE Trans. Wireless Commun., vol. 13, no. 12, pp. 6727–6740, Dec. 2014.

[84] C.-P. Chien, Y.-C. Chen, and H.-Y. Hsieh, "Exploiting spatial reuse gain through joint mode selection and resource allocation for underlay device to device com munications," in Proc. 15th Int. Symp. WPMC, Sep. 2012,

pp. 80–84.

[85] C.-H. Yu et al., "Resource sharing optimization for device-to-device communication underlaying cellular networks," IEEE Trans. Wireless Commun., vol. 10, no. 8, pp. 2752–2763, Aug. 2011.

[86] Lei Lei,Qingyun Hao,Zhangdui Zhong, "Mode Selection and Resource Allocation in Device-to-Device Communications With User Arrivals and Departures". IEEE Access 4: 5209-5222 (2016).

[87] Jun LI, G. L. Lei,, G. Manogaran . Shi, C.Wang, and G.Mastorakis, `` D2D Communication Mode Selection and Resource Optimization Algorithm With Optimal Throughput in 5G Network," IEEE Access, vol. 7, pp. 25263 _25273, Oct. 2018.

[88] W. Sun, E. G. Strom, F. Brannstrom, K. C. Sou, and Y. Sui, "Radio resource management for D2D-based V2V communication," IEEE Trans.Veh. Technol., vol. 65, no. 8, pp. 6636–6650, Aug. 2016.

[89] L. Liang, G. Y. Li, and W. Xu, "Resource allocation for D2D-enabled vehicular communications," IEEE Trans. Commun., vol. 65, no. 7, pp.3186–3197, Jul. 2017.

[90] Le Liang,Shijie Xie,Geoffrey YeLi: Graph-Based Radio Resource Management for Vehicular Networks, IEEE international conference on communication (ICC). 31 July 2018.

[91] Ahlem Masmoudi,Suhir Feki, Kais Mnif , and Faouzy Zarai : Efficient Radio Resource Management for D2D-based LTE-V2X communications.AICCSA 2018 : 1-6.

[92] Jie Mei, Kan Zheng, Long Zhao, Yong Teng, Xianbin Wang: A Latency and Reliability Guaranteed Resource Allocation Scheme for LTE V2V Communication Systems. IEEE Trans. Wireless Communications 17(6): 3850-3860 (2018).

[93] Giammarco Cecchini, Alessandro Bazzi, Barbara M. Masini, Alberto Zanella: Localization-based resource selection schemes for network-controlled LTE-V2V. ISWCS 2017: 396-401.

[94] Rongqing Zhang, Xiang Cheng, Liuqing Yang, Xia Shen, Bingli Jiao: A Novel Centralized TDMA-Based Scheduling Protocol for Vehicular Networks. IEEE Trans. Intelligent Transportation Systems 16(1): 411-416 (2015).

[95] Lianghai Ji, Man Liu, Andreas Weinand, Hans D. Schotten: Direct Vehicle-to-Vehicle Communication with Infrastructure Assistance in 5G Network. CoRR abs/1708.01020 (2017)

[96] Zhang Xiguang, Shang Yong, Li Xiaobo, Fang Jiayi ‘’Research on Overlay D2D Resource Scheduling Algorithms for V2V Broadcast Service’’. IEEE 84th Vehicular Technology Conference (VTC-Fall), Montreal, QC, Canada, Sep. 2016.

[97] Fakhar Abbas ; Pingzhi Fan ; Zahid Khan: “A Novel Low-Latency V2V Resource Allocation Scheme Based on Cellular V2X Communications”, IEEE Trans. Wireless Communications 18(7): 3850-3860.

[98] Luis F. Abanto-Leon, Arie Koppelaar, Sonia M. Heemstra de Groot: Graph-based resource allocation with conflict avoidance for V2V broadcast communications. IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC) PIMRC, 2017.

[99] M. I. Ashraf, M. Bennis, C. Perfecto, and W. Saad, “Dynamic proximity aware resource allocation in Vehicle-to-Vehicle (V2V) communications, ” in Proc. IEEE Globecom Workshops (GC Wkshps), Dec. 2016, pp. 1–6.

[100] Hao Ye and Geoffrey Ye Li .“ Deep Reinforcement Learning for Resource Allocation in V2V Communications,” IEEE Trans. on vehicular Technology., vol. 68, no. 4, pp. 3163 - 3173, Apr. 2019.

[101] Meiyan Wu, Yi Ren, Ping Wang, Chao Wang, Yusheng Ji , Location-Partition-Based Resource Allocation in D2D-Supported Vehicular Communication Networks. VTC Spring 2018: 1-6

[102] Molina-Masegosa, R.; Gozalvez, J. LTE-V for Sidelink 5G V2X Vehicular Communications: A New 5G Technology for Short-Range Vehicle-to-Everything Communications. IEEE Vehicular Technology Magazine, 2017, 12(4), pp. 30-39, doi:10.1109/MVT.2017.2752798.

[103] Giammarco Cecchini, Alessandro Bazzi, Barbara M. Masini, Alberto Zanella: MAP-RP: Map-based resource reselection procedure for autonomous LTE-V2V. PIMRC 2017: 1-6.

[104] Taylan Sahin, Mate Boban Radio Resource Allocation for Reliable Out-of-Coverage V2V Communications. VTC Spring 2018: 1-5.

[105] White Paper:” First Steps in 5G, Overcoming New Radio Device Design Challenges Series," Keysight Technologies’, 5992-2707EN, March, 2016. Accessed: Apr. 16, 2019. [Online]. Available: https:// literature.cdn.keysight.com/litweb/pdf/5992-2707EN.pdf

[106] Sidelink resource allocation mechanism for NR V2X, document R1-1902801, TSG RAN WG1 Meeting #96, 3GPP, Athens, Greece, Feb. 2019.