INTER-NUMEROLOGY INTERFERENCE MINIMIZATION IN 5G: A DEEP REINFORCEMENT LEARNING BASED APPROACH

by

Tuğrul Can Erk

B.S., Electrical and Electronics Engineering, Boğaziçi University, 2023

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Electrical and Electronics Engineering
Boğaziçi University
2023

ACKNOWLEDGEMENTS

First of all, I would like to express my sincere gratitude to my supervisor Prof. Ali Emre Pusane, my thesis supervisor, for his endless support, patience and understanding. I cannot express how grateful and lucky I am to be working with him and guided by his knowledge.

I am grateful to Prof. Tuna Tuğcu and Prof. Hacı İlhan for their participation in my thesis committee and for their precious feedback.

I would like to express my gratitude to my dearest mother, Nuran Erk, and my dearest father, Ömer Erk. I am deeply thankful to them for their support, and prayers, and for raising me to these days.

I would like to express my thanks to my brother, İsmet Erk, and Sevilay Erk, for their support in writing the thesis.

Lastly, I cannot find enough words to express how thankful I am to my dear wife, Melis Sena, for her trust and encouragement even in difficult times. This work would not be possible without her endless support and her love. Thank you for being with me every step of the way.

ABSTRACT

INTER-NUMEROLOGY INTERFERENCE MINIMIZATION IN 5G: A DEEP REINFORCEMENT LEARNING BASED APPROACH

The use of 5G technologies has become prevalent over the years with the increased use of mobile devices, video services, and more. At the same time, 5G technology brings lower latency, higher reliability, and higher throughput than 4G technology. creation of network slices via software-defined networks and network virtualization functions enables these improvements. A network slice provides flexibility to the network. Each slice can be dynamically configured within itself via SDN/NVF. The network meets diverse requirements of diverse services by creating a network slice.

At the higher level, NVF is responsible for creating and managing network slices. The reflection of a network slice on the physical layer is RAN slicing. With the numerology concept merging with 5G, the radio spectrum and corresponding resources become configurative in a manner of bandwidth, sub-carrier spacing, etc. However, the numerology solution has one drawback. Changing sub-carrier spacing in the resource grid destroys the orthogonality principle in the traditional OFDM signals. This leads the network to face new interference, inter-numerology interference.

The solutions that enable these improvements have also brought optimization problems that we have not faced In 4G networks. The main problem is about optimization since 5G technology requires heterogeneous signals for heterogeneous services. Therefore, the optimal allocation of limited resources in order to prevent this problem is the essential aim of this paper, with the minimization of inter-numerology interference and maximizing channel capacity.

ÖZET

5G AĞLARDA NUMEROLOJI KAYNAKLI KIRIŞIM MİNİMİZASYONU: DERİN TAKVİYELİ ÖĞRENME DAYALI YAKLAŞIM

Yıllar geçtikçe mobil cihazların, video hizmetlerinin ve daha fazlasının kullanımının artmasıyla 5G teknolojilerinin kullanımı yaygınlaştı. 5G teknolojisi, 4G teknolojisine kıyasla daha düşük gecikme süresi, daha yüksek güvenilirlik ve daha yüksek verim sağlar. Yazılım tanımlı ağlar ve ağ sanallaştırma fonksiyonları aracılığıyla ağ dilimlerinin oluşturulması, bu gelişmelere olanak sağlar. Her dilim, yazılım tanımlı ağlar ve ağ sanallaştırma fonksiyonları aracılığıyla kendi içinde dinamik olarak yapılandırılabilir. Ağ, bir ağ dilimi oluşturarak farklı hizmetlerin farklı gereksinimlerini karşılar.

Daha üst katmanda ağ sanallaştırma fonksiyonları, ağ dilimlerinin oluşturulmasından ve yönetilmesinden sorumludur. Bir ağ diliminin fiziksel katmana yansıması RAN dilimlemedir. Numeroloji konseptinin 5G ile birleşmesi ile radyo spektrumu ve ilgili kaynaklar, bant genişliği, alt taşıyıcı aralığı vb. açısından yapılandırılabilir hale gelmiştir. Numeroloji çözümünün bir dezavantajı vardır. Frekans ve zaman uzayında alt taşıyıcı aralığının değiştirilmesi, geleneksel OFDM sinyallerindeki diklik ilkesini yok eder. Bu, ağın numerolojiler arası girişim ile karşı karşıya kalmasına neden olur.

Bu iyileştirmeleri sağlayan çözümler, 4G ağlarında karşılaşmadığımız optimizasyon sorunlarını da beraberinde getirdi. 5G teknolojisi, heterojen hizmetlere heterojen sinyaller gerektirdiğinden asıl sorun optimizasyonla ilgilidir. Bu makalenin temel amacı, bu optimizasyon problemini çözmektir. Bunu yaparken de INI minimize etmeyi ve kanal kapasitesini en üst düzeyde kullanmayı amaçlar.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

a	Action
$e_m^u(k)$	CSI error of user u in MVNO m on subchannel k
$g_m^u(k)$	CSI of user u in MVNO m on subchannel k
$\hat{g}_m^u(k)$	Imperfect CSI of user u in MVNO m on subchannel k
$h_m^u(k)$	Small scale fading of user u in MVNO m on subchannel k
$I_m(k, k')$	INI power on subchannel k caused from subchannel k' for MVNO m
k	Subchannel
m	MVNO
$P_T(k)$	Transmission power on subchannel k
R	Reward
s	State
u	User
W	Bandwidth
α	Learning rate
α_m^u	Large scale fading of user u in MVNO m
β	Discount rate
$\gamma_m^u(k)$	Signal to interference plus noise ratio (SINR) for each user u from MVNO m using subchannel k
Δf	Subcarrier spacing
ϵ	Exploration rate
η_i	Spectrum usage ratio of numerology index i
θ	DNN weights
μ	Numerology index
$\sigma_{shadowing}$	Variance of shadowing
σ_w	Gaussian white noise

LIST OF ACRONYMS/ABBREVIATIONS

3GPP	3rd Generation Partnership Project
BS	Base Station
CP	Cyclic Prefix
CP-OFDM	Cyclic Prefix OFDM
CSI	Channel State Information
DNN	Deep Neural Networks
DQN	Deep Q Networks
DRL	Deep Reinforcement Learning
eMBB	Enhanced Mobile Broadband
FFT	Fast Fourier Transform
F-OFDM	Filtered OFDM
ICI	Inter Channel Interference
INI	Inter Numerology Interface
IoT	Internet of Things
ISI	Inter Slice Interference
LTE	Long Term Evolution
MDP	Markov Decision Process
MIMO	Multiple Input Multiple Output
ML	Machine Learning
mMTC	Massive Machine Type Communications
MVNO	Mobile Virtual Network Operators
NFV	Network Function Virtualization
NO	Network Operator
NSN	Narrow Subcarrier Spacing
OFDM	Orthogonal Frequency Domain Multiplexing
OoBE	Out of Band Emission
PAPR	Peak to Average Power Ratio
PRB	Physical Resource Blocks

QoS	Quality of Service
RAN	Radio Access Network
RE-LU	Rectifier Linear Unit
SCA	Successive Convex Approximation
SCS	Subcarrier Spacing
SDN	Software Defined Network
SINR	Signal to Interference plus Noise Ratio
SLA	Service Level Agreement
TI	Transmission Interval
URLLC	Ultra Reliable and Low Latency Communications
W-OFDM	Windowed OFDM
WSN	Wide Subcarrier Spacing

1. INTRODUCTION

With the arrival of 5G networks, it is expected to achieve higher data rates and throughput according to 4G networks. The improvement deployed with 5G networks can be interpreted in two categories, evolutionary and service-oriented. The former is focusing on performance improvements. It explains how 5G networks achieve 1000× traffic volume, 100× devices, and 100× throughput [1]. It gives a more structural explanation of the physical side of the 5G networks. In order to support 100× devices with 1000× traffic volume, novel technologies have been studied, such as multiple input multiple output (MIMO), millimeter wave communication, radio access network (RAN) slicing, spectrum efficiency, etc. [2].

The latter is concentrating on services. It investigated what kind of services can be supported by achieving a more reliable and faster communication infrastructure provided by 5G Networks. Nowadays, we are encountering different applications, such as massive usage of Internet of Things (IoT) devices and sensors that communicate with each other and are connected to the Internet in industries and homes, including remote-controlled devices and vehicles, remote surgeries, connected vehicles and vehicular networks, enhanced mobile communication, etc. Each service is merging with the innovation of the 5G network. Also, each service differs from each other according to their dependencies on reliability, throughput, and speed. Furthermore, with the transformation to 5G networks, it is expected to support these varying platforms and services and more upcoming applications.

Many factors, including but not limited to the exponential increase in the consumption of mobile video services and the use of IoT, have nudged institutions towards 5G systems [3]. In 5G multi-service networks, the three fundamental services include enhanced-mobile broadband (eMBB), ultra-reliable and low latency communications (URLLC), and massive machine type communications (mMTC), [4–6] which 4G Networks can not provide.

4G LTE networks are insufficient to cover these service requirements as they use one type of RAN. Unlike 4G networks, 5G technologies use RAN slicing to meet the different needs of different services. For instance, for URRLC, high reliability and low latency are crucial. For this reason, URRLC service uses a wider subcarrier spacing, whereas mMTC leans towards smaller subcarrier spacing [7]. RAN slicing enables us to have several independent virtual networks on the unique physical network infrastructure. Thus, a single physical network has to be divided into units of logical networks assigned to cover different services. This corresponds to varying configurative physical resource blocks in the RAN.

Software-defined network (SDN) and network function virtualization (NFV) are new merging techniques that allow network operators to deploy their heterogeneous services in one common infrastructure, [5, 8]. The former offers automation and dynamic reconfiguration of mobile networks while operating via defined software modules. The latter solves the challenges of using one common infrastructure by managing and providing flexibility to network functions. A network slice can be defined as the combination of these techniques. It is a set of network functions defined by software and tailored to guarantee users' and network operators' requirements, [9].

With SDN/NFV, a network slicing policy can be deployed in the one common physical infrastructure to meet service level agreement (SLA) and quality of services (QoS) requirements of different users and services, [10]. This new concept can compute network parameters and how to allocate resources, such as spectrum, power, etc., to each network operator. This refers to sharing and partitioning a physical network into several SDNs with various customized and optimized services and different configurations for the physical layer that diverge from traditional 4G networks.

In any 4G network, there is only one type of configuration in the RAN. The concept improvements made possible in 5G networks is called numerology. With the advancements in numerology, we can assign different physical resource blocks (PRB) to

different services. For example, larger resource blocks are deployed to URRLC services while narrower ones are assigned to mMTC [4].

While 4G physical layers use 15 kHz orthogonal frequency-division multiplexing (OFDM) blocks as PRB, 5G with numerology uses 15 kHz x 2μ OFDM signals as PRB, yielding different outcomes. This distinction allows flexibility in heterogeneous resource allocation, [11]. Nonetheless, this kind of varying OFDM signals leads to the loss of the orthogonality principle in OFDM. Due to this numerology concept, a new type of interference emerges in 5G networks, which is called INI (i.e., inter-numerology interference) or ISI (inter-slice interference).

There will be a degradation in system performance due to INI, which creates an optimization problem as the system should replace these resource blocks in the most efficient way. In past work, there are some approaches to solving this matter. Some of the researchers have tackled this problem using power allocation [12], while others have approached this as a problem of linear programming [13], successive convex approximation (SCA) with power constraint [14], or Langarian duality [15]. Recent literature has also focused on selecting the most efficient numerology for a set of services [16]. Along with these methods, there exist DRL applications to find optimal numerology selection or optimal resource allocation [17–20].

1.1. Contribution of Thesis

This thesis mainly benefits from [17] and builds upon the existing findings. This paper's significant contribution to the literature is to propose a way to minimize INI under imperfect channel state information (CSI) conditions, on top of using a binary non-convex optimization problem for the redistribution of resources. The main goal of [17] was to minimize the INI power while maximizing channel capacity. While this objective is a main concern of this paper as well, our methodology differentiates from theirs in terms of our approach to CSI. To put it simply, while the authors of [17] assume that the base station perfectly knows CSI, we train a deep reinforcement learning (DRL)

agent with Q-learning under the assumption that CSI is not known. The DRL agent is responsible for finding an optimal spectrum allocation that maximizes channel capacity and minimizes INI. With the method of Q-learning and the help of a deep neural network, the DRL agent finds a state-action pair that produces the maximum expected channel capacity for this non-convex optimization problem. The main contributions of this thesis can be listed below:

- Designing a binary non-convex optimization problem that maximizes channel capacity of users from different network slices for allocating physical radio spectrum of 5G RANs. While solving this optimization problem, the work accounts for the imperfect knowledge about channel state information, such as small-scale and large-scale fading. Moreover, the work also focuses on the interference that is mutually generated from the usage of multiplexing different OFDM signals, known as INI.
- the work presents a DRL model to find a solution to the optimization problem in order to reduce its computational complications. The designed DRL agent was trained to consider small-scale effects and INI power by multiplexing different numerologies in one common physical spectrum.
- The work also compares the achieved channel capacity of DRL-based resource allocation with the optimal allocation and static allocation. With DRL-based resource allocation, the DRL agent approximates the optimal allocation and performs better static allocation.

1.2. Outline of Thesis

This thesis is organized as follows: In Chapter 2, related works about network slicing, INI, and resource allocation are presented. The common idea of these works mainly depends on the question of how to deploy a network slice into the physical layer. The researchers studied this question with different aspects, from more homogeneous solutions to more heterogeneous solutions. In other words, they tried to find an an-

swer to how coarse the network configuration should be. Also, they implemented and analyzed different OFDM signals from more physical perspectives.

In Chapters 3 and 4, the system model of 5G RANs and the proposed solution of this thesis is explained. The problem of INI power and its minimization are studied depending on how a base station allocates its physical resource blocks and DRL optimization. In Chapter 5, the simulation results of the DRL method are presented. The conclusion part of the thesis is presented in Chapter 6.

2. RELATED WORKS

With the merging of new concepts in 5G networks, the researchers have been studying SDN/NFV-enabled networks, network slicing, spectrum isolation, new OFDM configurations, and new emerging interference analysis and its cancellation. In the literature, the challenges about allocation of network resources are taken into account with different solutions such as optimizing with respect to SLA and QoS constraints, training ML models to solve this optimization problem, configuring new OFDM signals, analyzing INI, and offering cancellation schemes of inferences.

The challenges of network slicing are that the common and scarce physical resources should distributed to each network operator, and each network slice should be managed individually. Also, these resources should tailored to grantee different constraints. In other words, the main challenge behind deploying a network slicing policy can be considered as an optimization problem.

The studies evaluate this optimization problem with different aspects. Some researchers convert tailoring network slices problem as matching problem. The system needs to match requirements with system capacity. Their studies answer the question of how to select the correct numerology and how to decide to deploy new network slices and configure these network slices. Furthermore, the other investigations are about applying ML techniques to find optimal solutions. In addition, the others try to solve this problem with a more physical approach. Their studies depend on optimally placing varying physical resource blocks into the resource grid. Also, some researchers analyzed the new inference caused by using different OFDM signals, INI. They inspected various types of OFDM signals and formulated the produced INI power of these OFDM signals. Moreover, they presented cancellation schemes for the INI interference between different network slices.

The studies tried to find a better way of improving the performance of 5G network slices and to find an answer to the questions listed below.

2.1. Slicing Configuration

The network owner makes available its resources to mobile virtual network operators, called tenants [21]. With network virtualization and software-defined networks in 5G, the tenants have some more specialized requests about network slices from the network owner. The requests of network slice may vary in different ways, such as bandwidth, data rate, latency, or reliability, in order to support different services. As a result of that, in 5G networks, each slice has a different configuration. However, creating a new slice increases system complexity in terms of increasing interference, computational complexity, and communication cost [22]. In [23, 24], the researcher listed the design challenges of network slicing.

The researcher in [25] developed a broker architecture that deals with tenant-network operator relations. This architecture was developed on the capacity broker defined by 3GPP for RAN sharing [26]. The developed 5G Network Slice Broker objective is to manage network resources and establish new network slices. The Broker monitors and forecasts traffic load. Also, it configures RAN scheduling. At the higher layer, the broker organizes inter-slice resource sharing. The broker configures the radio spectrum allocation policy at the lower layer for tenants' SLA needs.

In the study [27], the researcher focused on three key responsibilities of the network slice broker: traffic prediction per slice, making decisions for each tenant network slice request, and forecasting traffic load. They proposed a forecasting aware-network slicer algorithm to improve system utilization. The algorithm uses Holt-Winters theory for traffic forecasting and maps slice requests into geometric knapsack problem.

The researchers in [16], [28,29] approach this problem in a more physical manner. They consider this problem as how to select the proper numerology for each tenant.

Since a new numerology set can be considered as a new network slice, the network operator should select the proper numerology and slice configuration for each tenant's request. In [28], the researcher came up with a Mondrian forest model that takes service requirements, channel, and data traffic into account in order to optimize numerology selection for each sub-band. They focused on the statistical properties of a context vector that represent service requirements, channel, and data traffic. The proposed ML model achieves the QoS satisfaction level of %80-95.

The researcher of [29] labels each possible situation that can be present in RAN slicing. The situations are defined according to the number of numerology usage and guard bandwidth. They proposed an ML model that considers channel information and user information with service type. Then, the ML model decides on configurable parameters. The ML model classifies and labels input data according to each possible situation.

Moreover, another study from the researcher in [29] introduces a metric to optimize numerology selection, [16]. They explained a trade-off between numerology selection and flexibility in the system. Once a new numerology is presented, the system becomes more flexible and also more exposed to the degradation in performance in terms of interference and spectral efficiency. Furthermore, they have listed user and service requirements. According to these requirements, an optimization algorithm is proposed that decides the usage of guard band and numerology selection.

2.2. Optimization Problems

In 4G LTE networks, the physical layer supports only one type of OFDM signal. The resulting resource grid contains one type of physical resource block. However, the physical resource blocks and corresponding OFDM signals are changed with 5G Networks. The physical resource blocks can be different in size with respect to frequency and time. This led us to an allocation problem. Since the resource grid in the time and frequency domain is scarce, the network operator should place each physical resource

block into the grid according to users' different SLA and QoS requirements that are defined by different network slices. In the literature, the resource allocation problem is converted to an optimization problem that aims to place blocks with different sizes into the resource grid according to constraints on the SLA and QoS requirements.

In [15], the researchers proposed a solution to achieve higher spectral efficiency by using linear programming and Lagrangian duality. The NP-hardness of the 2-dimensional flexible resource allocation is solved with a scale-able optimization algorithm. They represent each requirement from different services as demand in throughput and latency tolerance. Although their results show great improvements, they did not consider physical effects and channel model such as INI, ICI.

The researchers in the [13, 30] formulated an optimization problem to achieve maximum data rates and considered the interference such as inter-symbol, inter-carrier, inter-band, and inter-numerology interference. In [13], they formulated the problem as an integer linear program and applied guard bands between slices to reduce interference. Their further study in [30] puts another constraint on latency requirement. The problem is converted to the max-min Knapsack problem.

In [14], the researchers focused more on SLA and QoS requirements of eMBB and URLLC services. They jointly allocate resources of multiple base stations to maximize eMBB users' throughput while considering short packet transmission of URLLC users and guaranteeing reliable communication for them. The optimization problem is formulated as a mixed-integer nonlinear programming problem. They proposed a low-complexity algorithm with successive convex approximation.

2.3. ML Techniques

The optimization problem mentioned above is a complex problem. The objective function can not represented in closed form [18]. In order to consider all constraints, such as latency requirements for URLLC users, high throughput requirements

for eMBB users, and support higher user density for mMTC users, the optimization problem and corresponding objective function need to be higher dimensional. This complexity can be reduced, which results in reducing representativeness of the 5G system model. Recently, this situation can be solved by ML techniques. An ML model can learn the system behavior without representing it in closed form.

For resource allocation and scheduling, the researchers in [18–20] proposed several deep reinforcement models. In [18], they developed a DRL model that optimizes mini-slot allocation for URLLC users over the time-transmission interval of eMMB users. The optimization problem in [18] is how many subcarriers should preempted for URLLC users from eMMB users so that loss in throughput of them is minimized. The method for scheduling is called punctured scheduling. The resulting model outperforms state-of-the-art algorithms proposed in the 3GPP standards body. However, the DRL model does not consider any channel model or INI while allocating subcarriers for URLLC users.

The research in [19] combines how to select network slice parameters such as numerology, bandwidth, and transmission power in order to isolate slices. They partitioned the resource grid into three parts, eMBB slice, URLLC slice, and shared slice. The shared slice allocates physical resource blocks according to punctured scheduling is used. The DRL model in the study aims to maximize aggregate throughput among all users. Also, the DRL model gets a penalty when the data rate requirement of eMMB users and the latency requirement of URLLC are not met.

Also, in [20], the researcher considered multiplexing URLLC and eMMB users with mMTC users. They proposed a DRL model for finding an optimal spectrum allocation policy that maximizes the long-term throughput. They implemented a channel model with large-scale and small-scale fading effects. Moreover, since combining three services led to a large action space (spectrum allocation policy) for DRL, they eliminated the number of actions by prioritizing delay constraints on the URLLC service and data-rate constraints on the eMBB service.

2.4. New OFDM Signal Techniques

In 5G Networks, the physical layer suffers because of the INI effect. The OFDM signals are traditionally used in previous mobile networks and offer an orthogonality to the network. On the other hand, this orthogonality is destroyed with a mixed-numerology scheme in 5G networks. Researchers studied well-known OFDM signals in order to choose which OFDM signals are more appropriate for the mixed-numerology concept, [7], [31–35].

In [31], the researchers focused on F-OFDM signals and analyzed interference for the OFDM waveform. The reason why they choose f-OFDM is to obtain better localization in the frequency domain, despite the fact that once the signal is filtered, the time localization of the signal will decrease. Also, intra-numerology interference between subcarriers and between the same numerology is analyzed because of the time dispersion caused by filtering. Moreover, the researchers in [31] formalized an optimization problem for power allocation in mixed-numerology systems and maximizing spectrum efficiency. The solution to the optimization problem manages to obtain maximum spectrum efficiency while allocating applicable transmission power to minimize interference caused by mixed numerology and filtering.

Another sub-band filtering technique is studied in [32]. The researchers analyzed a multi-service sub-band filtered multi-carrier system. The system partitions the radio spectrum into sub-bands for each service. In order to eliminate OoBE of the OFDM signals, the system puts guard bands between sub-band-parts. For maintaining spectral efficiency and reducing OoBE, the system filters each sub-band. Moreover, in [32], the researchers provide a channel equalization scheme that includes ICI, INI, timing offsets, and phase noise.

The researchers in [33,34] studied CP-OFDM since using CP allows time domain symbol alignment. Without losing time-synchronization and orthogonality, they analyzed INI produced between different numerologies for the conventional CP insertion

method. Moreover, they proposed a new way of inserting CP. The proposed method utilizes the difference in CP duration.

The other waveform candidate stated in [36] is W-OFDM. The difference between CP-OFDM and W-OFDM is the usage of the windowing function in order to smooth the transition edges of the transmit window. It provides several advantages, such as less complex implementation, robustness to time asynchronous, and no peak-to-average power ratio overhead. In [7], the researchers studied deeply the INI and how it is generated for W-OFDM systems. They also presented an INI cancellation algorithm called soft-output ordered successive interference cancellation algorithm (OSIC).

In order to compare the performance of 5G waveform candidates, F-OFDM, CP-OFDM, and W-OFDM, field tests were performed in [35]. The researchers built a test bed and collected measurements from the real environment. Moreover, they tried varying parameters, such as windowing size, filter order, CP length, etc. The results showed that the F-OFDM is better than the other regarding robustness to high signal-to-noise ratio spectral efficiency. Also, in some cases, the researcher stated that optimal spectrum allocation can be achieved with F-OFDM, which can remove guard bands.

3. 5G RAN NETWORKS

In this chapter, the requirements for new 5G RANs are listed. The concept of new waveform and numerology are explained in detail. Furthermore, three main multiplexing methods are analyzed. At the end of this chapter, the channel model used in this work, which also includes large-scale and small-scale fading fluctuations, is presented with its INI power formulation.

3.1. Numerology

It is decided that 3GPP moves with OFDM signals as in LTE. 3GPP standardizes the OFDM signals in the 5G physical layer in order to meet multiple service requirements. The numerology scheme is defined in the below table. Mixing numerologies in one common physical layer can be done in different approaches.

The usage of OFDM provides several advantages and is improvable, although some of them are destroyed with numerology concept. Firstly, OFDM signals have high spectral efficiency, which is important for higher data rates at lower frequencies. Moreover, OFDM signals have less complexity at the baseband. Hence, with the 5G network, the number of connected devices, low-powered IoT sensors, and low-cost base stations will be enormously increased, and the usage of basic baseband signals become more vital. Also, OFDM signals can be localized in the time domain easily, and using cyclic-prefix in the signals makes it more stable to timing synchronization error [2]. Moreover, the orthogonality between subcarriers decreases interference. Despite the advantages, OFDM signals have a high peak-to-average power ratio (PAPR). Higher PAPR will put more load on low-powered IoT devices and low-cost base stations, which must be power-efficient. The other disadvantage is out-of-band emission (OoBe) is high. Traditionally, this effect is eliminated because of the orthogonality principle. However, mixed numerology in the physical layer destroys this principle, [37]. The interference between mixed numerologies has more influence with high out-of-band-emission.

Table 3.1. Numerology options μ.

Parameters	Numerology Option μ				
	0	1	2	3	4
Subcarrier Spacing (kHz)	15	30	60	120	240
OFDM Symbol Duration (μs)	66.67	33.33	16.67	8.33	4.17
CP Duration (μs)	4.69	2.34	4.17 | 1.17	0.58	0.29
Slot Duration (ms)	1	0.5	0.25	0.125	0.0625

In 5G, the numerology index, μ, represents different physical resource blocks. 3GPP standardized PRB as shown in Table 3.1. The PRB structure used in LTE has been continued here as well. The LTE structure is used as a base for 5G PRB. The bandwidth of 5G PRB can have values of multiples of 15 kHz (i.e., like in LTE). The numerology index μ defines the subcarrier spacing (SCS) in PRB as $2^{\mu} \cdot 15$ kHz, where μ= {0, 1, 2, 3, 4}. Time slot interval of PRB is $T = 1/2^{\mu}$ ms [16, 38].

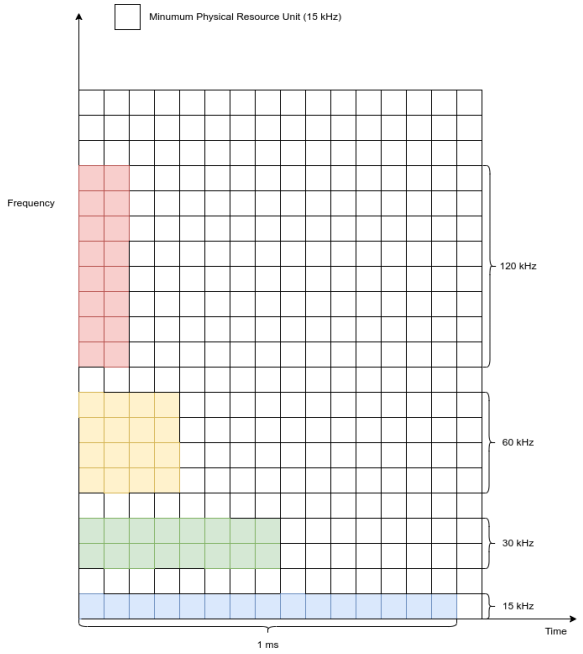


Figure 3.1. Representation of each PRB with different numerologies.

Each time-slot has 7 or 14 OFDM symbols with a length of time-slot interval T ms. The slots that use 7 OFDM symbols are called mini-slots and are presented to

meet latency-critical service requirements. In addition to that, each mini-slot contains 12 subcarriers with SCS of defined values above. The resulting PRB bandwidth can change from 180 kHz to 2880 kHz. The varying PRB in the resource grid can be seen in Figure 3.1, [30].

Table 3.2. 5G service requirements.

Service Type	Key Requirement	Spectral Efficiency	SCS	# of SC	Tcp	TI
eMBB	High throughput	High				
	High data rate	High				
mMTC	Support for data bursts		Large			
	High energy efficiency			Low		
URLLC	Low latency					Short
	High reliability		Large		Long	

In Table 3.2, the main 5G services are categorized according to their requirements. Table 3.2 summarizes which numerology can satisfy these requirements. According to [16], mMTC service should manage a large number of connected devices or IoT sensors with low power consumption. As a result of that, it is more appropriate to use lower numerologies for mMTC services. On the other hand, eMBB requires high spectral efficiency to achieve higher data rates and throughput. The intermediate numerologies are more suitable for eMBB service. Lastly, URLLC service depends on low latency with high reliability. To satisfy its need, higher numerologies should be selected.

3.2. Resource Allocation for Mixed Numerology

There are three ways of multiplexing different numerologies into the physical resource grid: time-domain multiplexing, frequency-domain multiplexing, and multiplexing in both domains. The method of multiplexing defines the flexibility of the system. When the multiplexing is done in only one domain, dedicated slices are created according to service needs. However, multiplexing in both domain support much more flexibility in the system. The system can assign any physical resource block to any

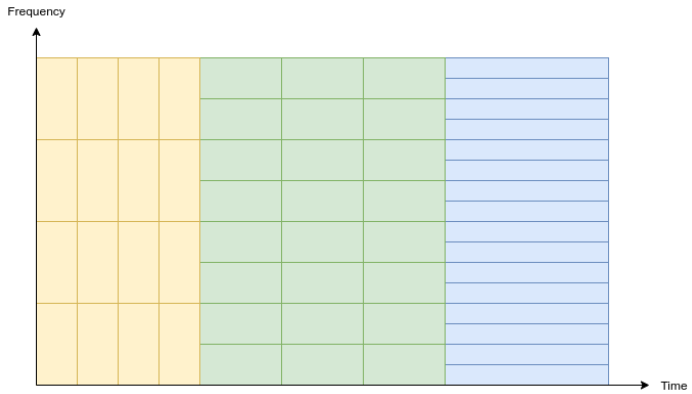
service. The trade-off of this flexibility is that the system must handle the complexity of resource allocation in an optimal way. The methods for multiplexing numerologies can be seen in the Figure 3.2.

3.2.1. Time-Domain Mixed Numerology

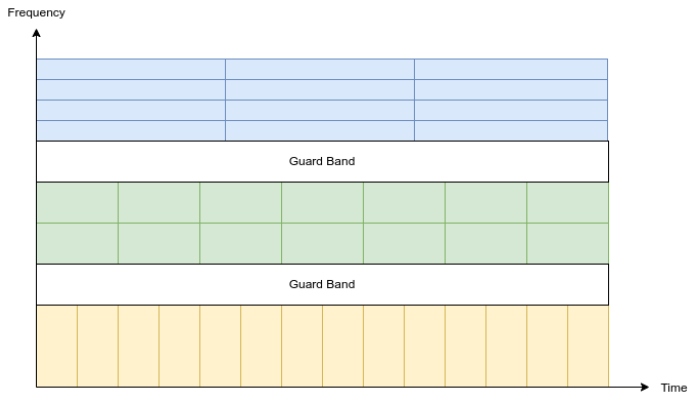
Using different cyclic prefixes for different subcarriers is one solution to multiplexing mixed numerologies in the time domain. With this scheme, every sub-frame is assigned to a different service. Within the subframe, only one numerology is used. As a result of it, INI is avoided and the radio spectrum usage becomes more effective since the whole sub-band uses the same OFDM configuration. However, dividing the resource grid by time domain may result in failing to meet time-critical service needs.

Using different SCS causes different varying lengths. For that reason, synchronization in the time domain is violated in the system. One way of synchronizing the system is to use varying cyclic prefixes. As shown in Table 3.1, SCS and CP duration from one numerology is an integer multiple of another SCS and CP duration from another numerology. This property helps synchronize OFDM symbols using the least common multiplier symbol duration, [33].

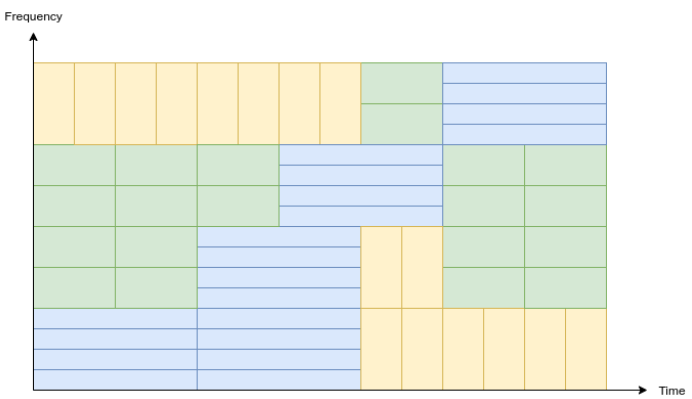
As shown in Figure 3.3 (a)-(b), [39], CPs are inserted into each symbol from different numerology, and synchronization in the time-domain is achieved. Also, by looking at their Fast-Fourier Transform window, it can be seen how INI affects each symbol in the time domain. In Figure 3.3 (c), the usage of common multipleir is represented. Common CPs are inserted at the beginning of the transmission time interval. As a result, symbol alignment is accomplished. Also, the INI minimization is better than individual CP insertion. In [33, 39], they proposed canceling this INI. [33], they proposed a new CP insertion method, represented in Figure 3.3.



(a)

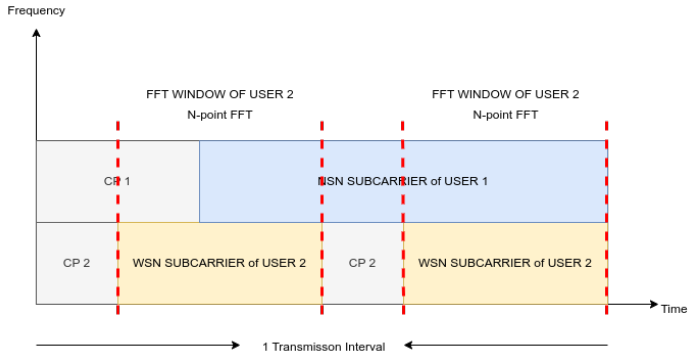


(b)

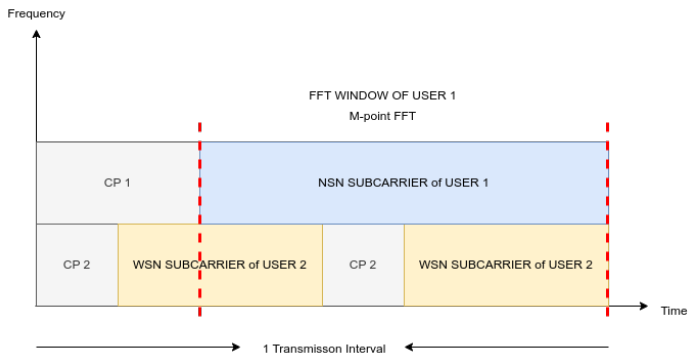


(c)

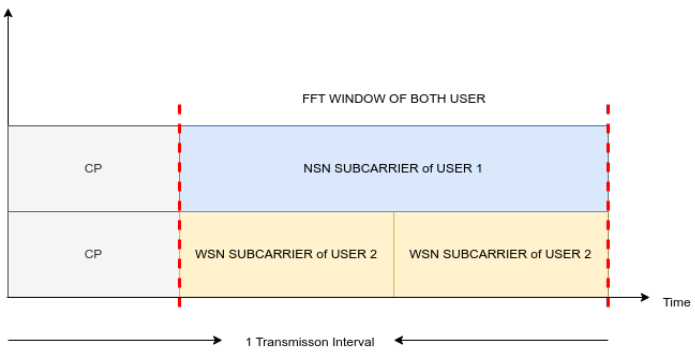
Figure 3.2. Representation of different multiplexing methods.(a)Time-domain multiplexing, (b)Frequency-domain multiplexing, (c)Time and Frequency multiplexing.



(a)



(b)



(c)

Figure 3.3. Time-domain numerology multiplexing representation of two users. NSN and WSN are assigned to User 1 and User 2, respectively. The FFT windows of each at the receiver side are represented in (a) and (b), where the symbol alignment is broken. For the common CP implementation in [33], (c) shows the symbol alignment in the time domain .

3.2.2. Frequency Domain Mixed Numerology

The network slices can be configured according to their spectrum needs. In one solution, the radio spectrum of 5G is partitioned into different sub-bands. Conceptually, each slice has its own spectrum. Moreover, by putting some guard bands, the out-of-band emission, which is a side effect of OFDM, is decreased. Also, this scheme is more suitable for supporting bursts of small data packages from URLLC services. However, applying some guard band between each network slice causes inefficient usage of the spectrum.

The OoBE can be decreased using F-ODFM [31] rather than using a guard band for the sake of spectrum efficiency. In the Figure 3.4, it can be seen that how OFDM side lobes will affect the other OFDM signal because of violating the orthogonality principle. Also, the Figure 3.4 shows how filtering can be applied to OFDM signals. However, the filtering process of the OFDM signal creates dispersion in the time domain.

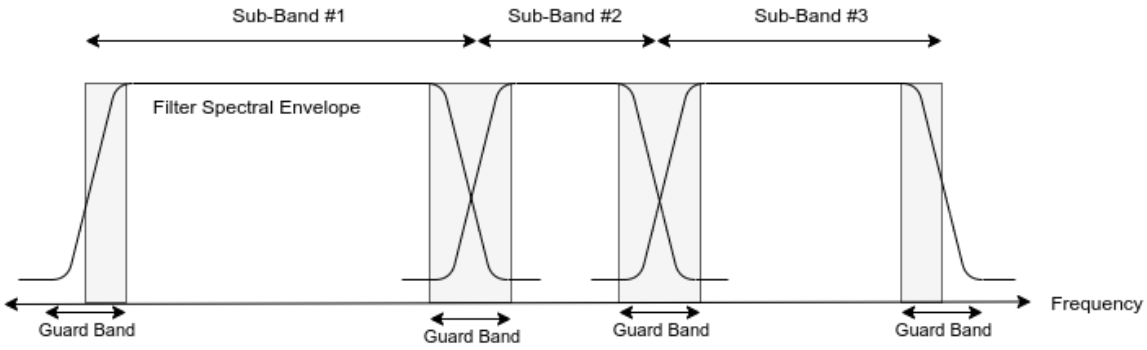


Figure 3.4. Frequency domain numerology multiplexing with guard bands.

3.2.3. Time and Frequency Domain Mixed Numerology

The most unrestricted solution for multiplexing mixed-numerology is utilizing the whole resource grid in both the time and frequency domain. By doing so, the usage of different sub-carriers in 5G is exploited.

The narrower sub-carrier is allocated into a mini-slot within the less delay tolerant Mobile band communication for delay-tolerant low-latency communication. However, this multiplexing scheme leads the system into a complex optimization problem. Although the system is more flexible, the network operator must find a way to allocate the spectrum in each slot. In each slot, 1 ms, the base station should consider the packet requests from all users from different services and try to assign time and frequency slots in an optimal way.

3.3. Channel Model

This work assumes that the network operator owns the base station providing radio interface to different mobile virtual network operators (MVNO) M. The NO allocates spectrum for each MVNO m. Each MVNO m is responsible for each packet scheduling. They maintain a logical radio interface and deliberate radio resources among its user U_m. Each m is responsible for meeting the SLA requirements of each of its users.

The NO scheduler uses a frequency-domain multiplexing approach. It divides the radio spectrum into K flat-fading subchannels. Each subchannel k can be thought of as creating a network slice at the physical layer. The network owner (NO) scheduler mainly uses a frequency-domain multiplexing approach. The NO assigns each subchannel k to each MVNO m. Each MVNO m uses different numerology as defined in Table 3.1.

We assumed that all users u estimate the channel state information (CSI) and report the channel gain to BS. However, this information is corrupted by some error. The CSI from each user u is imperfect. The channel gain is modeled by contributing large-scale fading, α_m^u, and small-scale fading $h_m^u(k)$. The large-scale fading consists of path-loos and shadowing effects. The shadowing is $N(0, \sigma^2_{shadowing})$. For the small-scale fading, the Rayleigh Fading model is used. So $h_m^u(k)$ is assumed as exponentially distributed with unit mean from the property of Rayleigh distribution and its power.

The formulation of subchannel gain is

$$g_m^u(k) = \alpha_m^u h_m^u(k), \tag{3.1}$$

where u_m represents the user u from MVNO m, and k corresponds to subchannel k.

For the imperfect CSI scenario, we added normally distributed Gaussian error to the subchannel gain. The formulation of subchannel gain under imperfect CSI is

$$\hat{g}_m^u(k) = g_m^u(k) + e_m^u(k), \tag{3.2}$$

where $\hat{g}_m^u(k)$ is the imperfect CSI and $e_m^u(k) \sim N(0, \sigma_e^2)$.

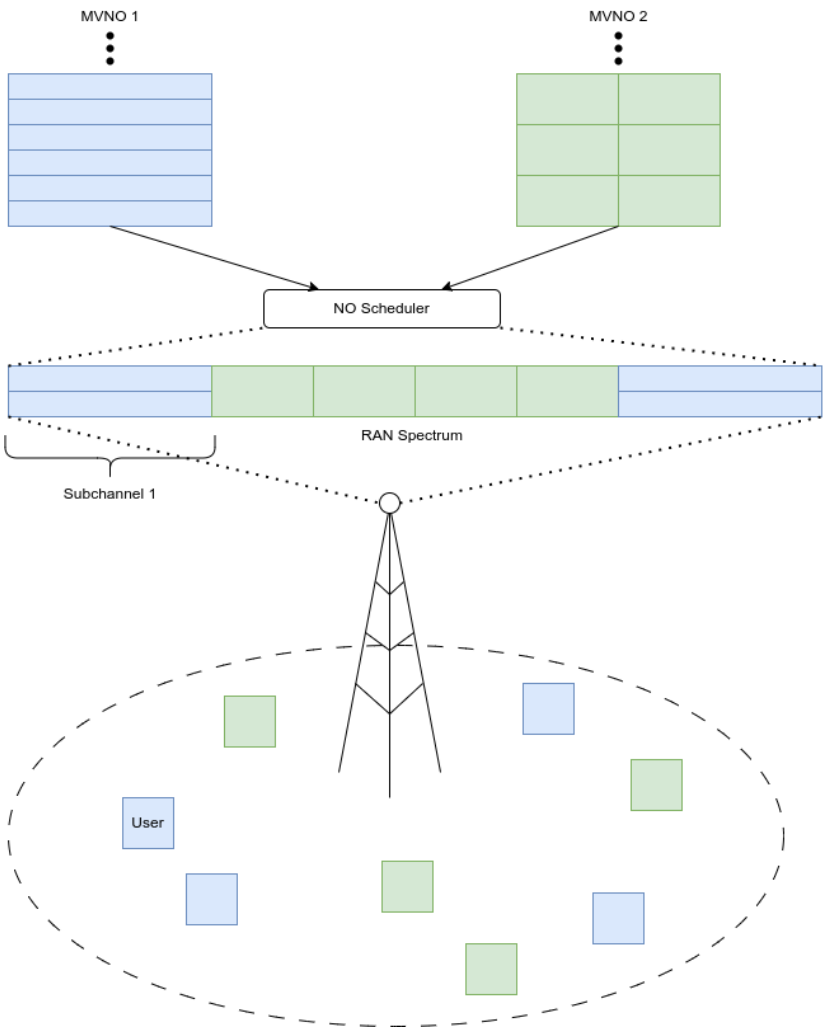


Figure 3.5. System model.

3.3.1. INI analysis

Since, in this multiplexing scheme and using mixed numerology, the system can no longer use the orthogonality principle of OFDM for the wireless communication channel. The OFDM signals from different subcarriers are overlapped, which causes INI. For a complete analysis of INI, the research from [33] is used.

For a multi-numerology system that has a radio spectrum of bandwidth W, it is considered that two users from different services using the spectrum with the ratio of η_1 and η_2, where $\eta_1 + \eta_2 = 1$. The users utilize num-1, num-2, and Δf_1, Δf_2 for their SCS and use individual CP configurations for the time alignment. Although the assumption consists of using two numerologies and a shared spectrum among them, any number of numerologies can be applied, and further analysis can be generalized.

We refer to two SCS, Δf_1, Δf_2, from num-1 and num-2 as narrow subcarrier spacing (NSN) and wider subcarrier spacing (WSN). throughout this chapter$Q$ represents $\Delta f_2/\Delta f_1$, which is an integer power of 2. N and M represent the FFT sizes of NSN and WSN, respectively. The relation is between N and M defined by Q, such that $N = QM$. The relation is also applied to CP sizes. The CP-sizes are N_{CP} and M_{CP}, respectively.

Once the spectrum is shared according to the ratios η_1 and η_2, we can express the subcarrier mapping for the spectrum and different numerology as

$$X_{nsn}(k) = \begin{cases} x_{nsn}(k), & 0 \leq k \leq \eta_1 N - 1 \\ 0, & \eta_1 N \leq k \leq N - 1 \end{cases} \tag{3.3}$$

and

$$X_{wsn}(l) = \begin{cases} 0, & 0 \leq l \leq (1 - \eta_2)M - 1 \\ x_{wsn}(k), & \eta_2 M \leq l \leq M - 1 \end{cases}. \tag{3.4}$$

The INI power expressions given in [33] are as follows,

- For INI power from k^{th} subcarrier of WSN to v^{th} subcarrier of NSN

$$I_{NSN}(k,v) = \frac{\rho^{WSN}(k)}{N \times M} \times \Psi(k,v), \tag{3.5}$$

for $0 \leq v \leq \eta_1 N - 1$ and $\{0 \leq k \leq \eta_2 N - 1 : k/Q \in \mathbb{Z}\}$, where

$$\Psi(k,v) = \frac{|sin[\frac{\pi}{Q}(1 + (1-Q)CP_R)(k-v)]|^2}{|sin[\frac{\pi}{N}(k - v + \eta_1 N)]|^2} + (Q-1) \times \frac{|sin[\frac{\pi}{Q}(1 + CP_R)(k-v)]|^2}{|sin[\frac{\pi}{N}(k - v + \eta_1 N)]|^2}, \tag{3.6}$$

$\rho^{WSN}(k) = |X_w ns(\eta_2 M + k/Q)|$ is the subcarrier power of k^{th} subcarrier of WSN, and CP_R is CP ratio such as $CP_R = N_{CP}/N = M_{CP}/M$. In order to consider the spectral distance between k^{th} and v^{th} subcarrier, the term $k - v + \eta_1 N$ is used. The k^{th} subcarrier of WSN translated to $k + \eta_1 N$.

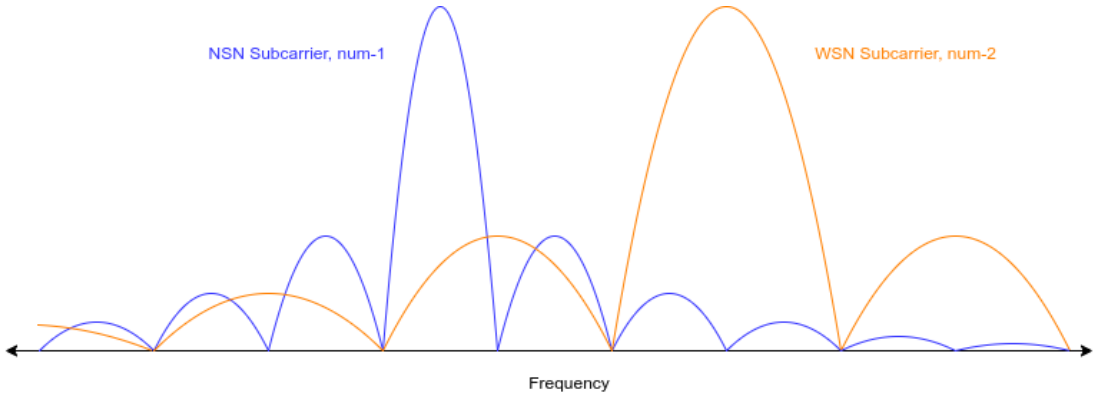
- For INI power from k^{th} subcarrier of NSN to p^{th} subcarrier of WSN

$$I_{WSN}(k,p) = \frac{\rho^{NSN}(k)}{N \times M} \times |\xi(k,p)|^2, \tag{3.7}$$

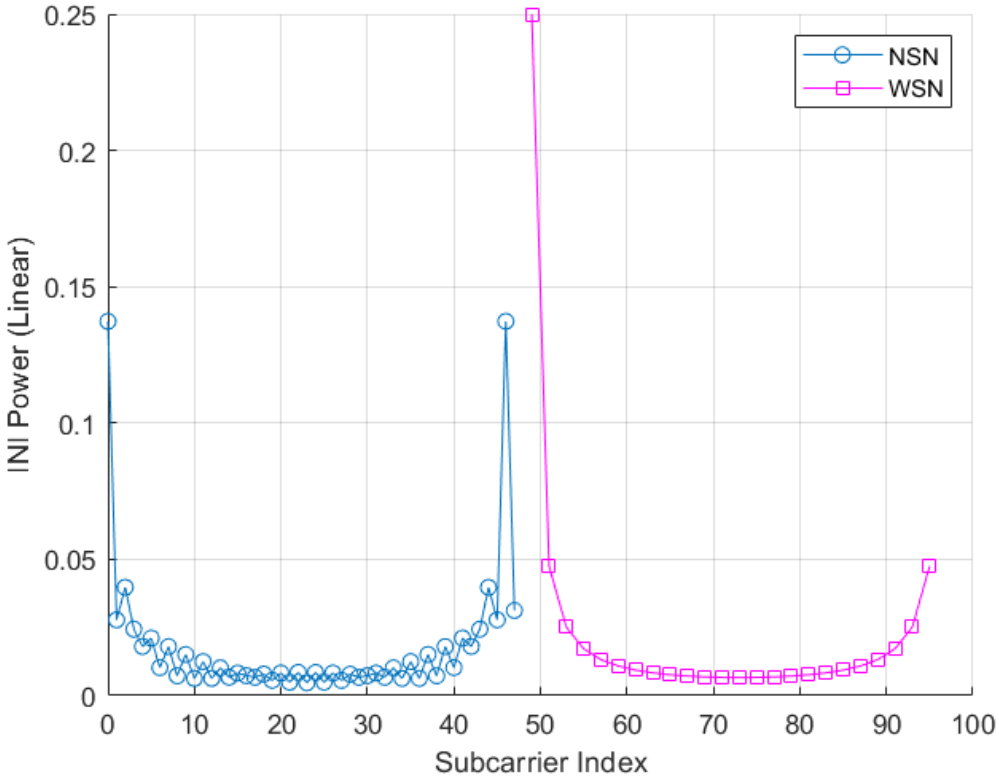
for $0 \leq k \leq \eta_1 N - 1$ and $\{0 \leq p \leq \eta_2 N - 1 : p/Q \in \mathbb{Z}\}$ where

$$\xi(k,p) = \frac{sin[\frac{\pi}{Q}(k-p)]}{sin[\frac{\pi}{N}(k - p - \eta_1 N)]}. \tag{3.8}$$

Remark 1: The term $sin[\frac{\pi}{Q}(k-p)]$ in $\xi(k,p)$ is zero for some values of k. This means that some subcarrier indices k of NSN do not interfere with WSN subcarriers. So, the mutually generated INI power is mostly from WSN and affects NSN subcarriers. The result can be seen in the Figure 3.6. The indices k actually correspond to the least common multiplier between two different numerologies.



(a)



(b)

Figure 3.6. (a) Two OFDM signals from different numerologies. The num-1 and num-2 signals are NSN and WSN signals, respectively. As can be seen, for every two OFDM lobes of the num-1 signal, the num-2 does not get any interference from num-1. (b) The INI power of num-1 and num-2. For the num-1, the INI power fluctuates.

The formulation of INI power on user $u \in U_m$ using subchannel k and due to subchannel k', by using Equation (3.6) and (3.8) is

- if the assigned numerology and corresponding SCS of subchannel k is smaller (INI power from WSN to NSN)

$$I_m^u(k,k') = \frac{P_T(k')}{M} \sum_{z=1}^{N} \sum_{v=1}^{M} \frac{g_u^m(k')}{N \times M} \Psi(z,v), \tag{3.9}$$

- if the assigned numerology and corresponding SCS of subchannel k is greater (INI power from NSN to WSN)

$$I_m^u(k,k') = \frac{P_T(k')}{N} \sum_{z=1}^{M} \sum_{v=1}^{N} \frac{g_u^m(k')}{N \times M} \xi(z,v). \tag{3.10}$$

The channel quality can be measured as the signal-to-interference-plus-noise ratio (SINR) for each user u from MVNO m using subchannel k. When INI and white Gaussian noise power, σ_w^2 , is added , the resulting SINR can be expressed as

$$\gamma_m^u(k) = \frac{P_T(k) g_u^m(k)}{\sigma_w^2 + \sum_{m' \neq m} \sum_{k' \neq k} x_{m',k'} I_m^u(k,k')}, \tag{3.11}$$

where

$$x_{m,k} = \begin{cases} 1, & \textit{if subchannel } k \textit{ is allocated for MVNO } m \\ 0, & \textit{otherwise} \end{cases}. \tag{3.12}$$

4. DRL DESIGN

Reinforcement Learning (RL) is a machine learning technique that aims to answer the question of how human interacts with their environment from a psychological point of view. The RL agent tries to formulate this interaction. States of the RL agent define the environment. In addition to that, an action space is defined for the RL agent. The RL agent interacts with the environment and observes its states, then takes an action from the action space. By taking action, the RL agent tries to reach a goal.

RL method is generally used for sequential decision-making. The RL agent makes trials and receives a corresponding reward or punishment. Through the learning process, the RL agent aims to get maximum cumulative reward and explores the environment through its trials. Since the learning procedure does not include prior knowledge about the environment, the RL agent does not need huge learning data sets. The RL agent will learn how to interact with the environment through its trials and explorations. This process is usually modeled as Markov Decision Process (MPD). The RL agent is in state $s \in S$, the set of states, and takes an action $a \in A$, the set of actions. This leads the RL agent to make a transition to the next state s' and to receive an intermediate reward R. MPD defines this transition by the probability, $P(s'|s,a)$. The cumulative discounted reward over time, G, is defined as

$$G_t = \sum_{i=0}^{t} \beta^i R_{t+i+}, 0 \leq \beta < 1, \tag{4.1}$$

where β is discount factor. It allows the RL agent to give more weight to the latest reward.

One of the main disadvantages of RL is once the state space is massive, it will be more challenging to find an optimum set of actions. The reason for that is that the RL agent may not observe every state-action pair. Also, the interaction between them may not be solved in a direct manner. To overcome this issue, a deep neural network is one possible solution.

To define a policy for each state-action pair, the Q-value function is used. $Q(s, a)$ takes state-action pairs and represents cumulative expected reward. As stated above, when we are faced with a large set of state-action pairs, the Q-value function becomes more complex and increases computational time. Using a DNN-based Q-value function (DQN) reduces the complexity of the Q-value function. The DNN approximates the actual Q-value function as $Q(s, a) \approx Q(s, a, \theta)$, where θ is DNN weights. During the training, the RL agent observes the environment and takes the action that produces the maximum reward. Then, according to the loss function, the DNN weights and are updated, and DQN approximates the actual Q-value function.

4.1. States

In this work, we are focusing on fading effects and INI power on each subchannel, $g_m^u(k)$ and $I_m^u(k, k')$ defined in Equation (3.2), (3.9), (3.10), respectively. Also, while contributing to these effects, we want to isolate each of them. In order to do that, $g_m^u(k)$ is set to 1 in Equation (3.9), (3.10). The set of states for the modeled environment, S, contains these effects and is defined as

$$S = \{G[k], I[k]\}_{k \in K}, \tag{4.2}$$

where the fading effect and INI power are defined as

$$G[k] = \{g_m^u(k)\}_{m \in M, u \in U_m}, \tag{4.3}$$

$$I[k] = \{ \sum_{m' \neq m} \sum_{k' \neq k} x_{m',k'} I_m(k, k') \}_{m \in M}. \tag{4.4}$$

Since the fading fluctuations are eliminated from INI power by setting $g_m^u(k) = 1$, the index u is dropped from the notation in Equation (4.4). As a result of that, we aimed for the DRL agent to learn better about INI.

4.2. Actions

We designed the DRL agent to find an optimal spectrum allocation among subchannels and MVNOs. In other words, the action space is composed of unique permutations of different spectrum assignments. The action space A can be defined as

$$A = \{X_1, X_2, ..X_N\}, \tag{4.5}$$

where N is equal to the number of all possible assignments and

$$X_i = \begin{bmatrix} x_{1,1} & ... & x_{1,K} \\ ... & \ddots & ... \\ x_{1,M} & ... & x_{M,K} \end{bmatrix}. \tag{4.6}$$

The X_i's are consist of $x_{m,k}$ in Equation (3.12).

4.3. Reward

Our aim for the DRL agent is to find optimum radio spectrum allocation. The quality of interaction with the environment in terms of facing fading fluctuations and INI power is calculated by the reward R at time t. For the formulation of the reward, the DRL agent should satisfy two requirements:

- The INI power on each subchannel k should be minimized.
- The highest channel capacity should be provided to each MVNO m.

The requirements and action space lead to the following optimization problem:

$$max \sum_{m=1}^{M} \sum_{k=1}^{K} \frac{1}{U_m} \sum_{u=1}^{U_m} x_{m,k} W log_2(1 + \gamma_m^u(k)), \tag{4.7}$$

subject to constraints on subchannel allocation, that can be written as

$$\sum_{k=1}^{K} x_{m,k} = S_m, \forall m \in M, \tag{4.8}$$

$$\sum_{m=1}^{M} x_{m,k} \leq 1, \forall k \in K, \tag{4.9}$$

where S_m represents the number of subchannels that are assigned to MVNO m. The reward is designed to satisfy these requirements and to solve the optimization in Equation (4.7). In order to improve the learning capability of the DRL agent, a high discount factor β is used. Thus, the DRL agent can focus more on the current state and expected reward at the present time slot.

Since we are considering imperfect knowledge about CSI, the channel outage probability is another consideration that DRL agent should seek. For this purpose, the DRL agent gets zero reward, or punishment, once it exceeds the actual channel capacity. The outage probability is modeled as a normal distribution for simplicity.

4.4. DRL Algorithm

We deploy DQN with the ϵ-greedy algorithm. The ϵ-greedy makes the DRL agent more flexible and allows it to explore the environment. The DRL agent takes action randomly with the probability of ϵ, otherwise, the action is chosen that maximizes the current policy π from the Q-value function. The optimal policy can be written as,

$$Q_\pi(s,a) = \max_{a \in A} Q(s,a), \tag{4.10}$$

where the Q-value function is

$$Q(s,a) = E[G_t | S_t = s, A_t = a]. \tag{4.11}$$

The Q-value function is updated every time step by DNN weights according to the loss function and learning rate. Thus, the Q-value function approximates its actual value and DNN weights are updated. For the update of the Q-value function, the formulation is

$$Q(s_t,a_t) \leftarrow (1-\alpha) Q_\pi(s_t,a_t) + \alpha[R(s_t,a_t) + \beta \max_{a_{t+1} \in A} Q(s_{t+1}, a_{t+1}; \theta)]. \tag{4.12}$$

The following loss function is minimized for updating DNN weights θ

$$L(\theta) = \sum_D [R(s_t,a_t) + \beta \max_{a_{t+1} \in A} Q(s_{t+1}, a_{t+1}; \theta) - Q_\pi(s_t, a_t; \theta)]^2, \tag{4.13}$$

where D represents the mini-batch that is sampled from the experience-replay buffer. The buffer stores $(s_t, a_t, s_{t+1}, a_{t+1})$ that generated throughout the training. In 4.1, the

training procedure of the DRL agent that gives the optimal spectrum allocation policy S_m is explained. The DRL algorithm is inspired by the research in [17].

Algorithm 1 DRL algorithm

for each episode **do**
  for each MVNO m $\in$ M **do**
    Randomly place U_m users in the BS coverage area;
    Calculate path-loss and shadowing effects for U_m users;
  end for
  for each time step **do**
    if time for fading update **then**
      Calculate small-scale fading for each subchannel k;
    end if
    Calculate INI;
    Calculate CSI reports;
    Observe the environment states s_t;
    Select the action a_t according to ϵ-greedy; {a_t implies the spectrum allocation policy}
    Store $(s_t, a_t, s_{t+1}, a_{t+1})$ in the experience-replay buffer;
    Uniformly sample a mini-batch size of D from the buffer;
    Update DNN weights θ by minimizing the loss function in eq. 4.13;
  end for
end for

Figure 4.1. Pseudo-Code for DRL algorithm.

5. SIMULATION & RESULTS

The simulation environment design with DRL Toolbox in MATLAB. The DRL agent is trained under these assumptions:

- We consider only one BS,
- The BS has 3 MVNO that numerology assignment varying from 15 kHz to 60 kHz,
- The BS distributes equal transmission power for all subchannels,
- Each MVNO serves the same number of users.

Under these assumptions, DRL is deployed with a fully-connected DNN with two hidden layers of 300 neurons each. These layers are activated with the Rectifier Linear Unit (Re-LU), $f(x) = max(0, x)$, function.

For minimizing the loss function in 4.13, RMSProp optimizer is chosen [40]. Also, Q-function values are updated with learning rate $\alpha = 0.001$. The configuration of DRL parameters is chosen according to the performance of varying hidden layers and DNN nodes. In order to avoid the local minima problem, the ϵ-greedy algorithm is implemented. As a result, the environment is explored randomly with a probability of ϵ. The exploration probability is decreased exponentially at the beginning of each episode.

Since it is assumed that the BS has imperfect knowledge about CSI, we applied an error to CSI. Furthermore, the imperfect CSI may cause channel outage. The outage probability means that the BS may exceed Shannon channel capacity, so the transmission may not occur. This probability is defined as 0.05 and 0.10.

Table 5.1. Simulation parameters.

Network Parameter	**Values**
Subchannel transmission power	18 dBm
Carrier frequency	2.5 GHz
Available numerologies	15, 30, 60 kHz
Subchannel bandwidth	720 kHz
Active users per MVNO	4
Fading statistic	Rayleigh
Fading update	1, 5 ms
CSI error variance	0.01, 0.05, 0.10
Outage probability	0.05, 0.10
Path loss model	36.7 log10 d + 33.05
Shadowing standard deviation	4 dB
Noise power	-115 dBm
DRL Parameter	**Values**
Learning rate	0.001
Discount factor	0.01
Experience-replay buffer size	1000
Mini-batch size	64
Episode duration	100 ms
Epsilon start	1
Epsilon stop	0.001
Exponential decaying factor	0.01

The training consists of 500 episodes that contain 100 time step. Each time step corresponds to 1 ms. Since users are placed at the beginning of each episode, the path-loss component does not change within episodes. In order to increase DNN performance, the states are isolated from this effect by normalizing the states with respect to path loss. The simulation parameters are listed in Table 5.1.

The performance of trained DRL agents is analyzed and compared with two different schemes, as listed below.

- Optimal Allocation: The optimal solution is searched in all possible allocation configurations for every time step,
- Static Allocation: The allocation scheme is chosen randomly from all possible allocations set at the beginning of episodes.

By comparing the DRL agent result with the former scheme, we can see how the DRL agent approaches to the optimum. With the latter scheme, it can be seen that the performance gain of the DRL agent.

Also, it is investigated that how the DRL agent responds to imperfect CSI, increasing the number of subchannels and the number of MVNOs. The first is related to the robustness of the DRL agent. Inserting an error and punishing the DRL Agent randomly according to outage probability shows how the system behaves when uncertainties about the channel increase. Moreover, by changing the fading update frequency, DRL agent behavior is analyzed against the uncertainties. For the second, the number of subchannels is increased from 3 to 6. In other words, we are increasing the number of action space for DRL agent. This means that the DRL agent will be challenged more to find the optimal solution. Lastly, by increasing the number of MVNOs from 2 to 3, we want to differentiate the dependency of INI on the number of subchannels or the number of multiplexed numerologies.

In Figure 5.1, the training result can be seen for different subchannel allocation scenarios. As expected, the training procedure is slowing down as the number of subchannels increases. For a lower number of subchannels, the DRL agent quickly learns the action space and finds optimal resource allocation easily. For simplicity, in Figure 5.1, only the first 200 episodes are shown instead of 500 episodes.

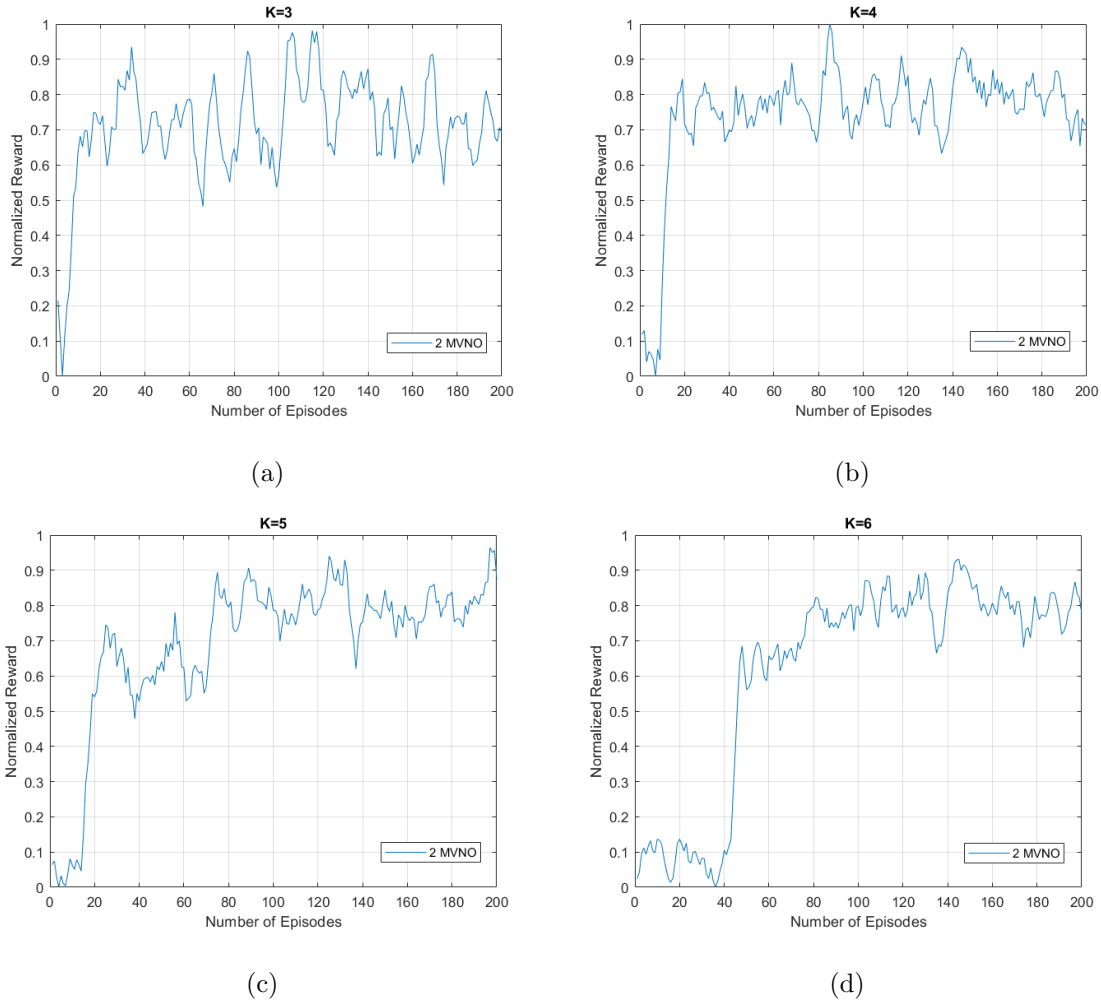


Figure 5.1. Normalized training curve of the DRL Agent when 2 MVNOs and the different number of subchannels, ranging from $k = 3$ to $k = 6$,(a)-(d).

The achieved channel capacity by the trained DRL agent is shown in Figure 5.2. In order to compare its performance, the results of optimal allocation and static allocation are added. The performance of the trained agent is approximating to the optimal allocation, even though the number of subchannel and multiplexed numerologies is increased.

Moreover, the Figure 5.3 shows that the performance loss in the 3 MVNOs scenarios. The reason behind that INI dynamics is more dependent on the number of multiplex numerologies rather than the number of subchannels used by MVNOs. The other reason is that enlarging the radio spectrum would cause decreasing INI as the distance between subchannels increases. As a result for that, increasing the number

of multiplex numerologies in the same spectrum causes performance loss more than expanding the radio spectrum while serving the same MVNOs.

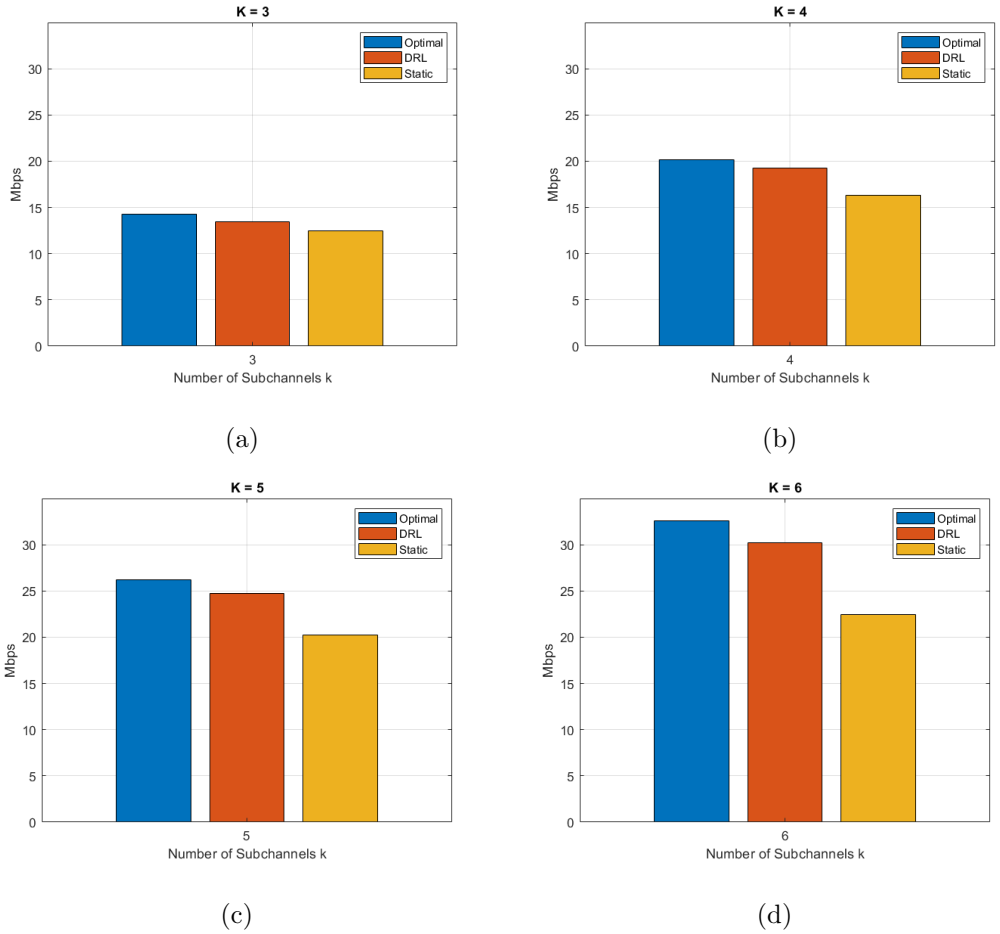


Figure 5.2. Achieved aggregate channel capacity with optimal allocation, trained DRL agent, and static allocation. The two MVNOs allocate the number of subchannels, ranging from $k = 3$ to $k = 6$,(a)-(d).

For imperfect CSI and channel outage scenarios, we analyzed the DRL agent in 2 different ways. The first is about changing CSI error variance and outage probability. For the first scenario, the fading update is 1 ms. The latter is about fading update frequency. The update frequencies are 1 and 5 ms. The CSI error variance σ and outage probability are assumed to be 0.01 and 0.10, respectively. The training and simulation results shown in the Figures 5.4, 5.5, 5.6 and 5.7. In all training and simulation below, the number of MVNOs M equals 2, and the number of subchannel K equals 3.

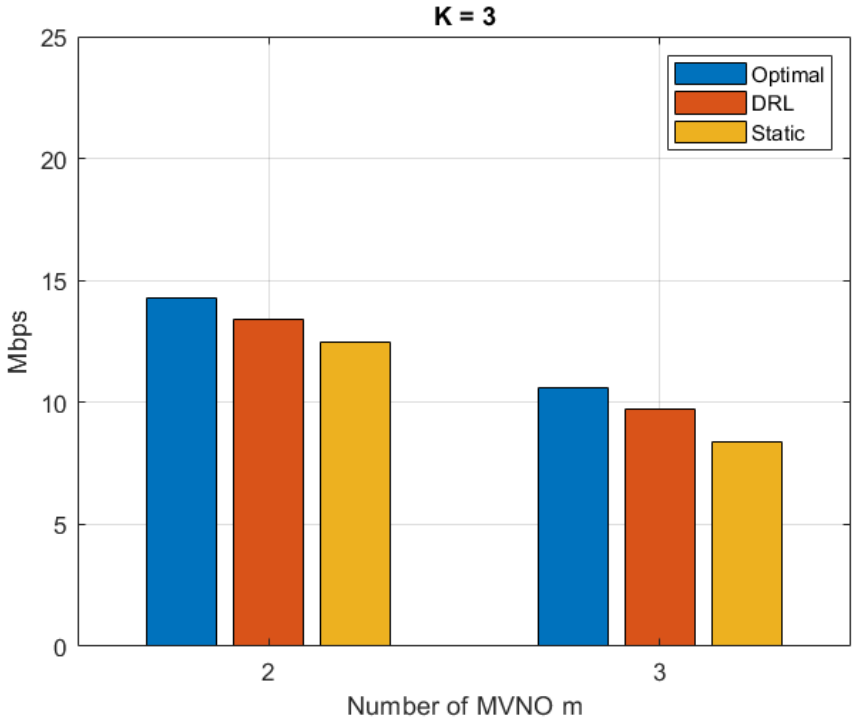


Figure 5.3. Achieved aggregate channel capacity of optimal allocation, trained DRL agent, and static allocation. The 2 and 3 MVNOs allocate the number of subchannels, ranging from $k = 3$ to $k = 6$. The number of multiplexed numerologies are $\{15\ kHz, 30\ kHz\}$ and $\{15\ kHz, 30\ kHz, 60\ kHz\}$, respectively.

In Figure 5.4, we consider imperfect channel condition. We introduce an error variance and corresponding outage probability. In order to see how the DRL agent is robust to channel fluctuations. As can be seen in the Figure 5.4, changing error variance has more impact on the training. The error variance slows down the learning process. Moreover, the fluctuations in the normalized reward are increased due to the outage probability, which is represented as a penalty for the DRL Agent.

In order to differentiate the effect of error variance and outage probability on the performance, the simulation results are collected in two ways. In the first way, the error variance σ_e is taken as constant, and outage probability varies. Conversely, the error variance σ_e values vary in the latter way. The fading update was equal to 1 ms through the simulations.

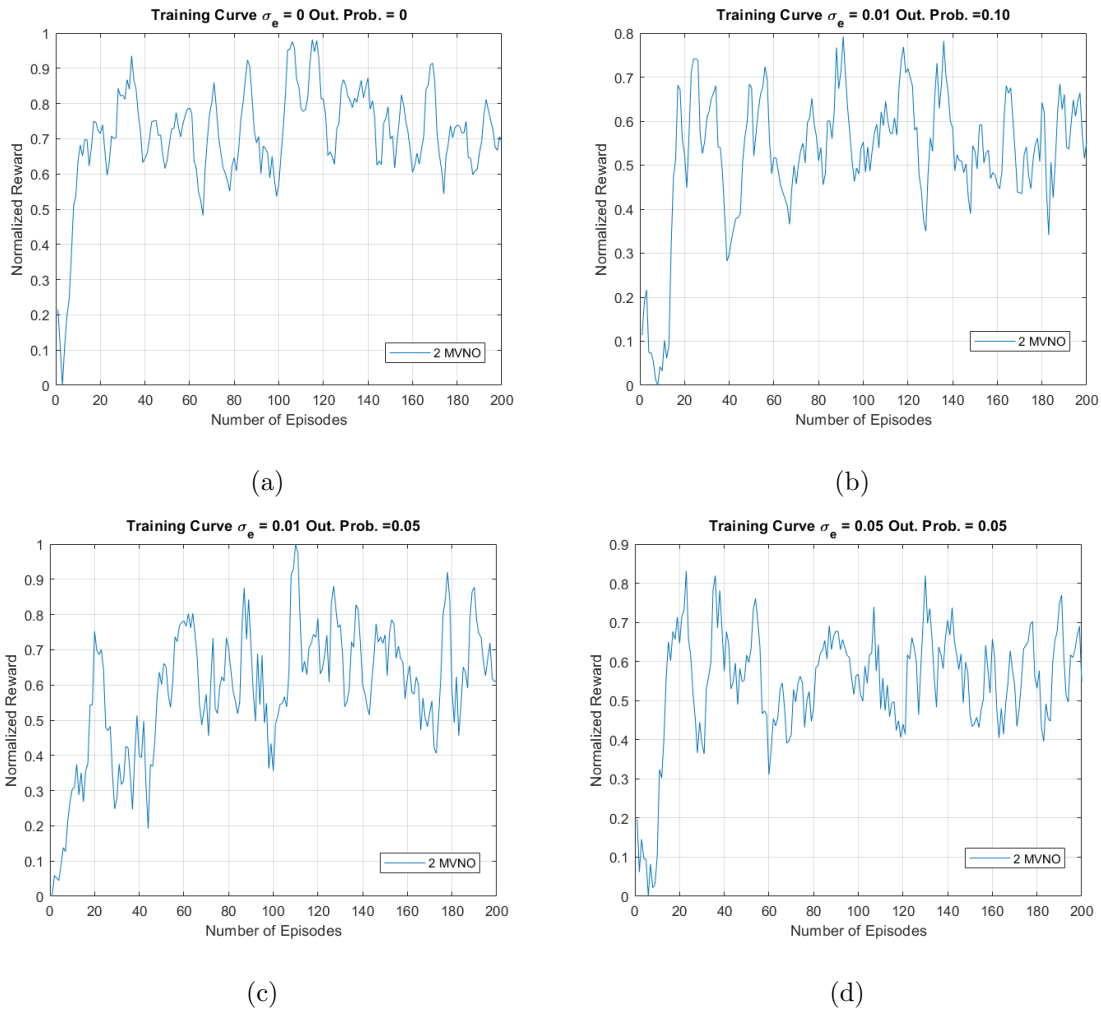


Figure 5.4. Normalized training curve of the DRL Agent when imperfect CSI condition is applied. The fading update is 1 ms. The imperfect CSI is implemented by changing error variance σ_e, 0 to 0.05, and outage probability, 0 to 0.10,(a)-(d).

In Figure 5.5, the outage probability takes values ranging from 0 to 0.10. Also, for the sake of measuring the performance loss, the simulation results of optimal and static allocation are presented. It is assumed that the CSI is perfectly known for optimal and static allocation. In this manner, we can see how much the DRL agent deviates from the optimal solution under imperfect CSI conditions.

As can be seen in the Figure 5.6, the effect of the error variance σ_e on the aggregated channel capacity is similar to the effect of the outage probability. The error variance σ_e takes values from 0 to 0.05. The optimal and static allocation, the perfect CSI case, is assumed, again.

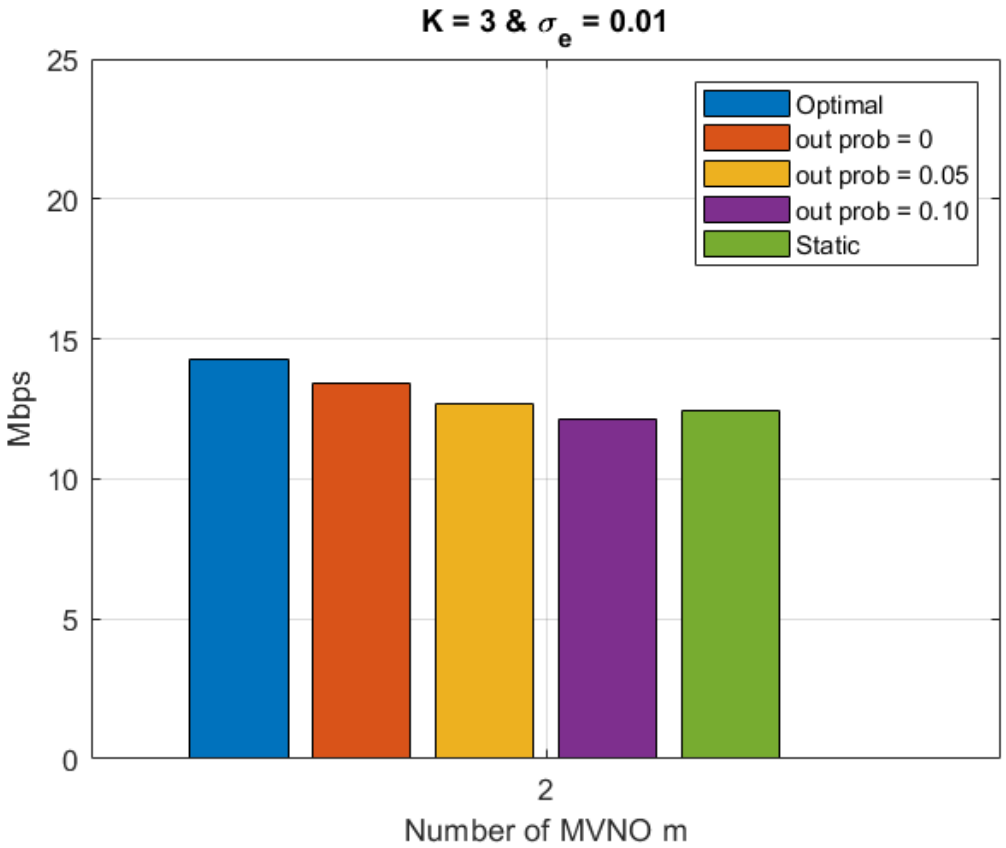


Figure 5.5. Achieved aggregated channel capacity of different outage probability values. The error variance σ_e value is 0.01. To compare performance, optimal and static with perfect CSI are presented.

In the Figures 5.5, 5.6, it is shown that the imperfect CSI situation can lead the DRL agent performs worse than static allocation. The performance loss is similar to static allocation according to optimal allocation. The edge case occurs when the error variance σ_e equals 0.01 and the outage probability equals 0.10. The reason for this performance loss is the fluctuations in the CSI and, consequently, in the channel capacity. The DRL agent struggles to figure out the system's behavior. The fluctuations in the reward can be seen in the Figure 5.4 (b).

To improve DRL performance when imperfect CSI occurs, we changed the fading update frequency from 1 ms to 5 ms. This results in a slowdown in small-scale fading and fluctuations in imperfect CSI. The reason behind that is to reduce the complexity that the DRL agent deals with. By slowing down the CSI update without losing the

concept of small-scale fading, the performance of the DRL agent increases, as can be seen in the Figure 5.7. The edge case mentioned above is also presented with optimal and static allocation to compare the performance gain by changing the fading update frequency.

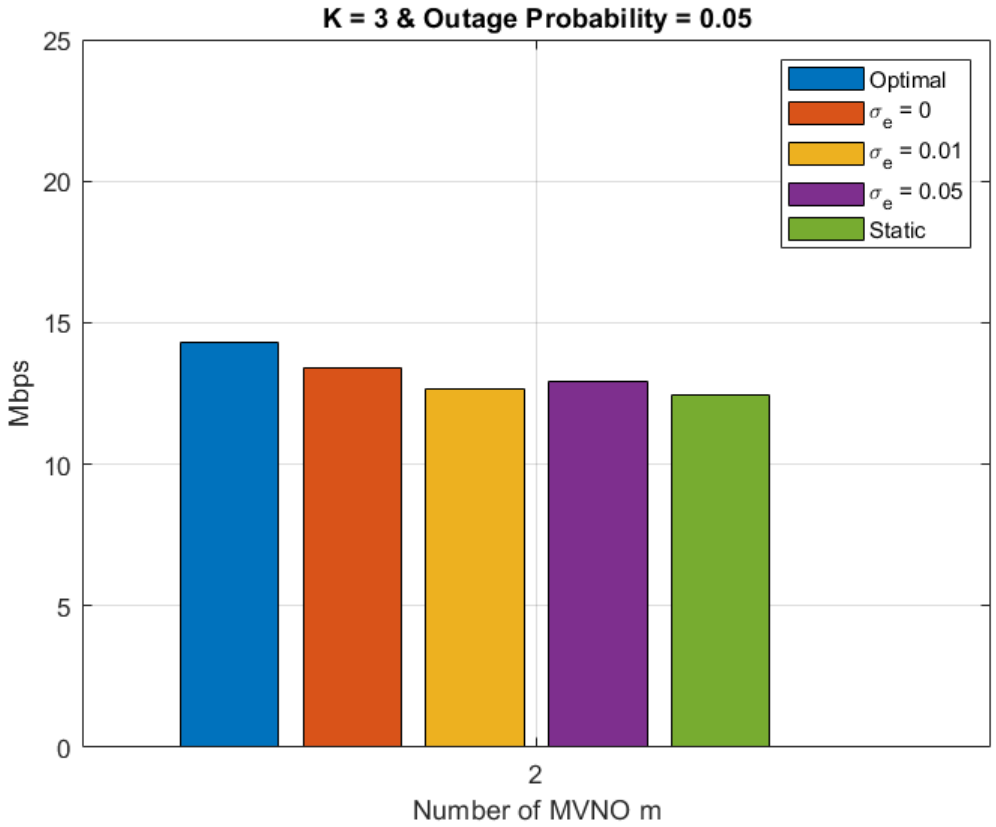


Figure 5.6. Achieved aggregated channel capacity of different error variance σ_e values. The outage probability value is 0.05. To compare performance, optimal and static with perfect CSI is presented.

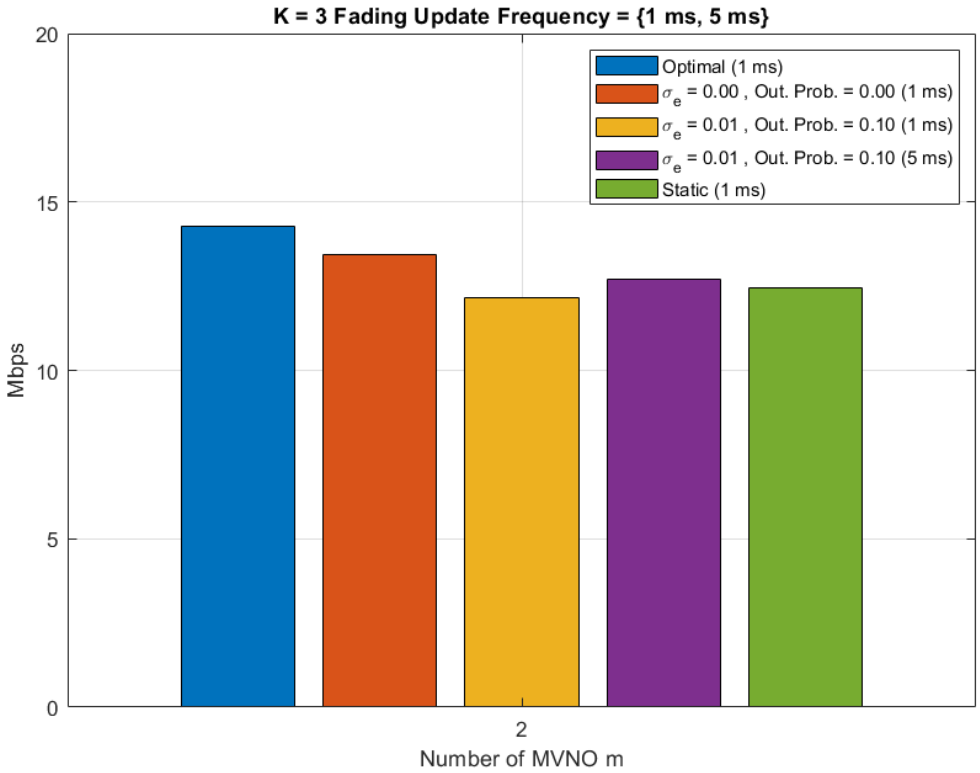


Figure 5.7. Achieved aggregated channel capacity for different CSI updates.

6. CONCLUSION

With 5G technology, the paradigm changes from homogeneous services and communication infrastructure to more heterogeneous and tailored services and resource allocation. Traditional mobile communication misses the concept of allowing users to cover their special needs. It offers one fit-for-all solution for every user and service [6, 41]. The homogeneous system in 4G LTE becomes more restricted for mobile network operators.

Although handling one fit-for-all solution traditionally offers some level of standardization among different users and lessens the burden on managing of scarce physical resources for the network operator. 5G networks bring more gains than its losses. 5G networks are more suitable for the customization of users' and network operators' needs. It is designed to tailor scarce physical resources to guarantee each needs. Furthermore, it provides a platform that supports varying multi-services. It allows remote surgeries by providing ultra-reliable and ultra-low communication, massive IoT devices, and connected sensor by handling a significant amount of data traffic volume. This allowance comes from the concept of network slicing and mixing numerology in the physical layer. However, one drawback of presenting a custom solution for each requirement is that it brings heterogeneity into the physical layer. By changing the structure of the RAN from traditional RANs, 5G RANs can use different OFDM that mutually generate interference between network slices.

In this work, we formalized the optimal physical resource allocation problem using mixed numerology. The work converts this optimization problem to a binary non-convex optimization problem while considering small-scale effects and INI. In order to solve this complex optimization problem, a DRL agent is trained to find optimal allocation and maximize channel capacity.

The work shows that the proposed DRL model is a good way of reducing computational complexity to solve resource allocation under the constraints of INI and fading effects while supervising channel capacity. The trained DRL agent overcomes the static allocation scheme in terms of performance.

Furthermore, the results show that when action space is relatively small, the DRL agent quickly learns the environment by associating INI power and small-scale fading with channel capacity. Also, the DRL agent outputs approach to the optimal allocation, although we introduce imperfection in CSI and outage probability.

REFERENCES

1. Foukas, X., G. Patounas, A. Elmokashfi and M. K. Marina, "Network Slicing in 5G: Survey and Challenges", *IEEE Communications Magazine*, Vol. 55, No. 5, pp. 94–100, 2017.

2. Zaidi, A. A., R. Baldemair, H. Tullberg, H. Bjorkegren, L. Sundstrom, J. Medbo, C. Kilinc and I. Da Silva, "Waveform and Numerology to Support 5G Services and Requirements", *IEEE Communications Magazine*, Vol. 54, No. 11, pp. 90–98, 2016.

3. Barakabitze, A. A., M. Liyanage and A. Hines, "QoESoft: QoE Management Architecture for Softwarized 5G Networks", *2020 IEEE International Conference on Communications Workshops (ICC Workshops)*, Dublin, pp. 1–6, 2020.

4. Zhang, L., A. Ijaz, J. Mao, P. Xiao and R. Tafazolli, "Multi-Service Signal Multiplexing and Isolation for Physical-Layer Network Slicing (PNS)", *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*, Toronto, pp. 1–6, 2017.

5. Yousaf, F. Z., M. Bredel, S. Schaller and F. Schneider, "NFV and SDN—Key Technology Enablers for 5G Networks", *IEEE Journal on Selected Areas in Communications*, Vol. 35, No. 11, pp. 2468–2478, 2017.

6. Kaloxylos, A., "A Survey and an Analysis of Network Slicing in 5G Networks", *IEEE Communications Standards Magazine*, Vol. 2, No. 1, pp. 60–65, 2018.

7. Zhang, X., L. Zhang, P. Xiao, D. Ma, J. Wei and Y. Xin, "Mixed Numerologies Interference Analysis and Inter-Numerology Interference Cancellation for Windowed OFDM Systems", *IEEE Transactions on Vehicular Technology*, Vol. 67, No. 8, pp. 7047–7061, 2018.

8. 3GPP, *System Architecture for the 5G System (Release 15)*, Technical Specification (TS 23.501), 3rd Generation Partnership Project (3GPP), 2017.

9. Sciancalepore, V., M. Di Renzo and X. Costa-Perez, "STORNS: Stochastic Radio Access Network Slicing", *ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*, Shanghai, pp. 1–7, 2019.

10. Alkhafaji, A. R. and F. S. Al-Turaihi, "Multi-Layer Network Slicing and Resource Allocation Scheme for Traffic-aware QoS ensured SDN/NFV-5G Network", *2021 1st Babylon International Conference on Information Technology and Science (BICITS)*, Babil, pp. 327–331, 2021.

11. Zambianco, M. and G. Verticale, "Interference Minimization in 5G Physical-Layer Network Slicing", *IEEE Transactions on Communications*, Vol. 68, No. 7, pp. 4554–4564, 2020.

12. Korrai, P. K., E. Lagunas, A. Bandi, S. K. Sharma and S. Chatzinotas, "Joint Power and Resource Block Allocation for Mixed-Numerology-Based 5G Downlink Under Imperfect CSI", *IEEE Open Journal of the Communications Society*, Vol. 1, pp. 1583–1601, 2020.

13. Marijanovic, L., S. Schwarz and M. Rupp, "A Novel Optimization Method for Resource Allocation Based on Mixed Numerology", *ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*, Shanghai, pp. 1–6, 2019.

14. Wang, Z. and V. W. Wong, "Joint Resource Block Allocation and Beamforming with Mixed-Numerology for eMBB and URLLC Use Cases", *2021 IEEE Global Communications Conference (GLOBECOM)*, Madrid, pp. 1–6, 2021.

15. You, L., Q. Liao, N. Pappas and D. Yuan, "Resource Optimization With Flexible Numerology and Frame Structure for Heterogeneous Services", *IEEE Communications Letters*, Vol. 22, No. 12, pp. 2579–2582, 2018.

16. Yazar, A. and H. Arslan, "A Flexibility Metric and Optimization Methods for Mixed Numerologies in 5G and Beyond", *IEEE Access*, Vol. 6, pp. 3755–3764, 2018.

17. Zambianco, M. and G. Verticale, "Spectrum Allocation for Network Slices with Inter-Numerology Interference using Deep Reinforcement Learning", *2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications*, London, pp. 1–7, 2020.

18. Huang, Y., S. Li, C. Li, Y. T. Hou and W. Lou, "A Deep-Reinforcement-Learning-Based Approach to Dynamic eMBB/URLLC Multiplexing in 5G NR", *IEEE Internet of Things Journal*, Vol. 7, No. 7, pp. 6439–6456, 2020.

19. Setayesh, M., S. Bahrami and V. W. Wong, "Resource Slicing for eMBB and URLLC Services in Radio Access Network Using Hierarchical Deep Learning", *IEEE Transactions on Wireless Communications*, Vol. 21, No. 11, pp. 8950–8966, 2022.

20. Suh, K., S. Kim, Y. Ahn, S. Kim, H. Ju and B. Shim, "Deep Reinforcement Learning-Based Network Slicing for Beyond 5G", *IEEE Access*, Vol. 10, pp. 7384–7395, 2022.

21. Tseliou, G., K. Samdanis, F. Adelantado, X. C. Pérez and C. Verikoukis, "A capacity broker architecture and framework for multi-tenant support in LTE-A networks", *2016 IEEE International Conference on Communications (ICC)*, Kuala Lumpur, pp. 1–6, 2016.

22. Abdellatif, A. A., A. Mohamed, A. Erbad and M. Guizani, "Dynamic Network Slicing and Resource Allocation for 5G-and-Beyond Networks", *2022 IEEE Wireless Communications and Networking Conference (WCNC)*, Austin, TX, USA, pp. 262–267, 2022.

23. Rost, P., C. Mannweiler, D. S. Michalopoulos, C. Sartori, V. Sciancalepore, N. Sastry, O. Holland, S. Tayade, B. Han, D. Bega, D. Aziz and H. Bakker, "Network Slicing to Enable Scalability and Flexibility in 5G Mobile Networks", *IEEE Communications Magazine*, Vol. 55, No. 5, pp. 72–79, 2017.

24. Ordonez-Lucena, J., P. Ameigeiras, D. Lopez, J. J. Ramos-Munoz, J. Lorca and J. Folgueira, "Network Slicing for 5G with SDN/NFV: Concepts, Architectures, and Challenges", *IEEE Communications Magazine*, Vol. 55, No. 5, pp. 80–87, 2017.

25. Samdanis, K., X. Costa-Perez and V. Sciancalepore, "From network sharing to multi-tenancy: The 5G network slice broker", *IEEE Communications Magazine*, Vol. 54, No. 7, pp. 32–39, 2016.

26. 3GPP, *Study on Radio Access Network (RAN) Sharing enhancements (Release 13)*, Technical Specification (TR 22.852), 3rd Generation Partnership Project (3GPP), 2014.

27. Sciancalepore, V., K. Samdanis, X. Costa-Perez, D. Bega, M. Gramaglia and A. Banchs, "Mobile traffic forecasting for maximizing 5G network slicing resource utilization", *IEEE INFOCOM 2017 - IEEE Conference on Computer Communications*, Atlanta, GA, USA, pp. 1–9, 2017.

28. Kotagiri, D., A. Sawabe, E. Takahashi, T. Iwai, T. Onishi and Y. Nishikawa, "Context-based Mixed-Numerology Profile Selection for 5G and Beyond", *2022 IEEE 19th Annual Consumer Communications & Networking Conference (CCNC)*, Las Vegas, NV, USA, pp. 611–616, 2022.

29. Yazar, A. and H. Arslan, "Selection of Waveform Parameters Using Machine Learning for 5G and Beyond", *2019 IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Istanbul, Turkey, pp. 1–6, 2019.

30. Marijanovic, L., S. Schwarz and M. Rupp, "Multi-User Resource Allocation for Low Latency Communications Based on Mixed Numerology", *2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall)*, Honolulu, HI, USA, pp. 1–7, 2019.

31. Mao, J., L. Zhang, P. Xiao and K. Nikitopoulos, "Interference Analysis and Power Allocation in the Presence of Mixed Numerologies", *IEEE Transactions on Wireless Communications*, Vol. 19, No. 8, pp. 5188–5203, 2020.

32. Zhang, L., A. Ijaz, P. Xiao, A. Quddus and R. Tafazolli, "Subband Filtered Multi-Carrier Systems for Multi-Service Wireless Communications", *IEEE Transactions on Wireless Communications*, Vol. 16, No. 3, pp. 1893–1907, 2017.

33. Kihero, A. B., M. S. J. Solaija and H. Arslan, "Inter-Numerology Interference for Beyond 5G", *IEEE Access*, Vol. 7, pp. 146512–146523, 2019.

34. Kihero, A. B., M. S. J. Solaija, A. Yazar and H. Arslan, "Inter-Numerology Interference Analysis for 5G and Beyond", *2018 IEEE Globecom Workshops (GC Wkshps)*, Abu Dhabi, pp. 1–6, 2018.

35. Guan, P., D. Wu, T. Tian, J. Zhou, X. Zhang, L. Gu, A. Benjebbour, M. Iwabuchi and Y. Kishiyama, "5G Field Trials: OFDM-Based Waveforms and Mixed Numerologies", *IEEE Journal on Selected Areas in Communications*, Vol. 35, No. 6, pp. 1234–1243, 2017.

36. Qualcomm Inc., *Waveform Candidates*, Standard Contribution (R1-162199), 3rd Generation Partnership Project (3GPP), Busan, Korea, Apr. 2016.

37. Mao, J., L. Zhang, S. Mcwade, H. Chen and P. Xiao, "Characterizing Inter-Numerology Interference in Mixed-Numerology OFDM Systems", ArXiv: Signal Processing arXiv:2009.13348, 2020.

38. 3GPP, *Technical Specification Group Radio Access Network; NR; Physical Layer; General Description (Release 15) V1.0.0*, Technical Specification (TS 38.201), 3rd Generation Partnership Project (3GPP), 2017.

39. Cheng, X., R. Zayani, H. Shaiek and D. Roviras, "Inter-Numerology Interference Analysis and Cancellation for Massive MIMO-OFDM Downlink Systems", *IEEE Access*, Vol. 7, pp. 177164–177176, 2019.

40. Ruder, S., "An overview of gradient descent optimization algorithms", ArXiv: Machine Learning arXiv:1609.04747, 2016.

41. Ferrus, R., O. Sallent, J. Perez-Romero and R. Agusti, "On 5G Radio Access Network Slicing: Radio Interface Protocol Features and Configuration", *IEEE Communications Magazine*, Vol. 56, No. 5, pp. 184–192, 2018.