

Article

Solution for Interference in Hotspot Scenarios Applying Q-Learning on FFR-Based ICIC Techniques

Iago Diógenes do Rego *  and Vicente A. de Sousa, Jr. Department of Communications Engineering, Federal University of Rio Grande do Norte,
Natal 59078-970, Brazil; vicente.sousa@ufrn.br

* Correspondence: iago.diogenes.072@ufrn.edu.br

Abstract: This work explores interference coordination techniques (inter-cell interference coordination, ICIC) based on fractional frequency reuse (FFR) as a solution for a multi-cellular scenario with user concentration varying over time. Initially, we present the problem of high user concentration along with their consequences. Next, the use of multiple-input multiple-output (MIMO) and small cells are discussed as classic solutions to the problem, leading to the introduction of fractional frequency reuse and existing ICIC techniques that use FFR. An exploratory analysis is presented in order to demonstrate the effectiveness of ICIC techniques in reducing co-channel interference, as well as to compare different techniques. A statistical study was conducted using one of the techniques from the first analysis in order to identify which of its parameters are relevant to the system performance. Additionally, another study is presented to highlight the impact of high user concentration in the proposed scenario. Because of the dynamic aspect of the system, this work proposes a solution based on machine learning. It consists of changing the ICIC parameters automatically to maintain the best possible signal-to-interference-plus-noise ratio (SINR) in a scenario with hotspots appearing over time. All investigations are based on ns-3 simulator prototyping. The results show that the proposed Q-Learning algorithm increases the average SINR from all users and hotspot users when compared with a scenario without Q-Learning. The SINR from hotspot users is increased by 11.2% in the worst case scenario and by 180% in the best case.

Keywords: ICIC; FFR; hotspot; ns-3; Q-Learning; machine learning



Citation: Diógenes do Rego, I.; de Sousa, V.A., Jr. Solution for Interference in Hotspot Scenarios Applying Q-Learning on FFR-Based ICIC Techniques. *Sensors* **2021**, *21*, 7899. <https://doi.org/10.3390/s21237899>

Academic Editor: Yang Yue

Received: 16 September 2021

Accepted: 26 October 2021

Published: 27 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

According to a Cisco forecast [1], by 2022, traffic from wireless and mobile devices will account for 71% of global IP traffic. Between 2017 and 2022, all data traffic in mobile networks will be seven times bigger, as a result of the intense sharing and consumption of data, especially video. Recent studies [2] predict that, over the next 7 years, 1.4 billion people will start using mobile internet for the first time. Until 2025, over 60% of the world's population will be using mobile internet. Additionally, there is an ongoing change of consuming habits. As more people consume video-related content, longer and more frequently, especially via mobile devices, the use of mobile data per user will be five times bigger by 2024.

Nonetheless, smartphones are not the only devices responsible for the increase on traffic demand. According to the GSMA Association [2], between 2018 and 2025, the amount of IoT devices in mobile networks will triple, reaching 25 billion connections. Furthermore, 5G will also increase the number of connected devices and use cases offered by the network. This constant growth of traffic demand calls for higher network capacity. In fact, the Shannon–Hartley theorem [3] demonstrates that increasing the available bandwidth is the most effective way to increase a channel's capacity. However, the spectrum is a limited resource that has to be used efficiently.

In order to increase network capacity, the evolution of wireless communication technologies has introduced approaches, such as the use of multiple antennas, power control,

management of co-channel interference by reusing the spectrum, and the implementation of more efficient modulation schemes. For example, Long Term Evolution (LTE) originally proposes a reuse factor of 1 in order to increase spectrum efficiency. A reuse factor of 1 means that all network cells operate in the same frequency band. A reuse factor of 3, on the other hand, decreases the band available to each cell by 3, decreasing the channel's capacity. However, a reuse factor of 1 may lead to low Signal-to-interference-plus-noise Ratio (SINR), especially to users located further away from the e-NodeB (eNB) (radio base station on LTE), due to interference from neighboring cells.

These techniques aim at increasing channel capacity, but they also introduce new challenges, mainly related to the compromise between coverage and quality of service. For example, enhanced Mobile Broadband (eMBB) has been a widely discussed topic in the literature. It represents the evolution of mobile communications and it has been the focus in the first commercial deployments of 5G. One of its goals is to provide high data rates and wide coverage. However, as 5G operates in high frequencies, the coverage tends to decrease, as the signal attenuation is more severe. Consequently, the network will likely have to deploy more base stations, increasing the need to mitigate interference from adjacent cells.

In addition, some aspects that have strong impacts on mobile network performances, such as the traffic demand or the amount and distribution of users, are often unpredictable or variant through time. Big cities usually host events that can last from a few hours to a couple of days and these can be either recurrent or a one time occasion. Events and locations, such as school fairs, football games, music festivals, school parades, food parks, and malls can suddenly increase the number and density of users at that location.

The planning and deployment of a mobile network is traditionally associated with modeling the channel based on the specific demand for a location. This usually does not take into account some parameters that can vary in an unexpected way, as the ones mentioned above. In non-dynamic systems, the occasional appearance of areas with high user concentration can severely degrade the channel quality and the system's ability to serve its users.

In this work, the term hotspot is used to describe this region with a high density of users and it does not mean a new Access Point (AP). Besides the increase on traffic demand, due to a high number of users, there is also a higher probability of increased interference, especially on systems with a reuse factor of 1. In such scenarios, the users in the borders of the coverage area are the most affected by interference.

The paper is organized as follows. Section 2 presents the problem of high user concentration and its impact on system performance, along with some classical solutions for this problem. Sequentially, Section 3 introduces Fractional Frequency Reuse (FFR), an important component of the proposed solution, and Section 4 shows how 3rd Generation Partnership Project (3GPP) has developed Inter-Cell Interference Coordination (ICIC) techniques over the years. Section 5 presents relevant related works and Section 6 indicates how the system is modeled for simulations. Finally, Sections 7–9 show results from preliminary analysis and Section 10 presents our proposed solution, as well as the proof-of-concept results and discussion.

In this paper, our key contributions are:

- Performance comparison of classic FFR-based ICIC algorithms that take into account the number of users and their distributions along the cell (edge and center);
- Statistical analysis of the strict frequency reuse scheme that identifies which parameters do not need to be dynamically controlled;
- Analysis of the strict frequency reuse algorithm's performance in a mixed scenario with hotspot and homogeneous user distribution;
- Q-Learning algorithm that continually operates in the network to dynamically mitigate the performance loss (SINR) that results from the appearance of hotspots (densely populated areas).

2. Interference in Hotspot Scenarios

Mobile networks face many challenges that can limit channel capacity. For example, the lack of available bandwidth is a common issue that can be addressed by increasing the spectrum reuse, as a way to increase the capacity limit. However, that may lead to a higher co-channel interference, which happens when more than one user is using the same radio resource at the same time.

The solution presented in this paper can be applied on an Orthogonal Frequency-Division Multiple Access (OFDMA)-based system (e.g., 4G LTE or 5G NR) in a Downlink (DL) direction, whose users may experience interference from multiple neighboring cells reusing the same frequency at the same time. This interference may be a severe limiting factor in terms of capacity, especially with a reuse factor of 1.

In such scenarios, users from different cells can interfere with each other, and users located in the cell edge are the most affected by co-channel interference [4]. These users are far from the transmitting cell and closer to adjacent cells, increasing the probability of receiving any interfering signals with higher power levels and signals from its serving cell with lower power levels. Hence, user distribution is another aspect that can have a negative impact on system performance.

Therefore, the existence of regions with high user concentration (hotspots), especially if close to the cell edge, can lead to a lack of available resources and a higher probability of users being allocated to the same frequency. Regardless of its location, the appearance of hotspots may increase interference, leading to lower SINR. This results in higher Bit Error Rate (BER) and, consequently, lower traffic capacity.

All of these challenges are well known and can degrade performance; however, they are partially predictable. On the other hand, some modern, urban hotspot scenarios bear a more dynamic aspect. For example, the number of people flowing through a big city does not merely repeat weekly, since extraordinary situations that are hard to predict often happen in large urban areas.

For instance, a certain neighborhood that consistently had a uniform distribution of users may change significantly because a new food park has opened, a soccer league changed its venue, a protest is taking place or an accident stopped the traffic. In such situations, mobile networks that were planned and deployed statically will not be able to serve its users well. Besides, the traffic demand per user is also likely to increase as these situations or events induce a different behavior in users. It leads them to consume more bandwidth by sharing photos and videos, aggravating the high traffic demand.

2.1. Classic Solutions for Interference in Hotspot Scenarios

2.1.1. MIMO Systems

Multiple-Input Multiple-Output (MIMO) is a technology that uses multiple antennas for transmitting and receiving signals to mitigate negative aspects of the channel and/or multiplex data transmission. The first goal, for example, can be attained by increasing the diversity of transmitted signals [5]. The unpredictability of the wireless channel is used as a tool to improve system performance [6] by using multiple antennas to exploit the advantages of spatial diversity [7]. Furthermore, MIMO can also explore spatial multiplexing, using multiple antennas to transmit data through various channels.

Given a minimum spacing between antennas, it is possible to obtain independent or weakly correlated channels on each Tx/Rx (transmission and reception) pair. This concurrent transmission in the same frequency band can increase data rates or improve reliability without compromising spectral efficiency.

A scenario with hotspots is prone to have some channels suffering with less interference than others. Thus, if an appropriate combination method is applied for the different transmissions, it is possible to improve performance. For example, choosing the path with better SINR [5] is a simple solution that can lead to better results. Additionally, multiple antennas can also control the signal amplitude and phase to direct beams towards certain users or regions and avoid users suffering from interference [8].

Recent wireless communication systems point MIMO as an efficient approach to increase data rates, especially when suffering from small-scale multi-path, such as systems with hotspots. For instance, the use of multiple antennas has been widely discussed in the literature as a core feature of 5G systems [9], through massive MIMO. The use of Millimeter Waves (mmWaves) enables antenna arrays with a high number of elements [10]. However, MIMO is a solution implemented in the physical layer, which requires investments in hardware and it often demands channel estimation.

Nevertheless, this work does not intend to indicate MIMO as a technique to be replaced. Still, it introduces FFR as an appropriate technology to address the specific problems noted in the last section. In mobile networks, MIMO often co-exists with FFR techniques.

2.1.2. Small Cells

Heterogeneous Network (HetNet) is an extended concept of Hierarchical cell structure (HCS) that has been discussed even before the standardization of LTE [11]. It is usually associated with urban scenarios, where the traffic demand and the number of users are constantly increasing.

In HetNets, different categories of cells coexist with each other, and they usually differ in coverage and capacity [12]. Still, they do not necessarily need to share the same radio access technology. In such deployments, small cells are an alternative to offload the macrocells traffic, especially on hotspots and cell-edge [13], due to severe interference.

Dense urban areas that already have limited capacity may require big investments to deploy another conventional access point. In such cases, the use of small cells is attractive, since they can improve coverage and transmit with lower power, given the smaller distance between user and transmitter. These cells usually operate in the same frequency band as the macrocells, transmitting with lower power to reduce interference.

Recent studies involving HetNets usually focus on two categories of small cells: femtocells and picocells [14–16]. The former is typically deployed by the user, it is not planned, and it may be private, being called a Closed Subscriber Group (CSG). The latter has the functionality of a regular eNB, but with smaller coverage and lower transmission power. Picocells are usually deployed to cover areas with a high density of users and they can also be deployed indoors, if necessary.

The use of small cells is also indicated as a key technology for 5G. The densification of the network has been indicated in the literature as an efficient solution to help satisfy the requirements for 5G [17].

However, the use of small cells also imposes some challenges. Its deployment demands investment in planning and hardware, and it may increase the amount of unnecessary hand-offs [18]. The large power difference between cell tiers can leave small cell users at a disadvantage. On the other hand, CSGs can create holes in coverage for the macrocell users, as they do not have access to the CSG.

3. Fractional Frequency Reuse

A mobile network usually offers coverage to its subscribers with cells that have a limited frequency band for transmitting and receiving signals. In order to use the spectrum efficiently, a reuse factor of 1 is commonly applied, which means that all of the cells use the entire available bandwidth.

In order to mitigate inter-cell interference, some networks increase the reuse factor. For example, a reuse factor of 3 creates a topology where adjacent cells do not share bandwidth, as illustrated in Figure 1. However, there is a significant loss in spectral efficiency, given that only 1/3 of the original bandwidth is available to each cell.

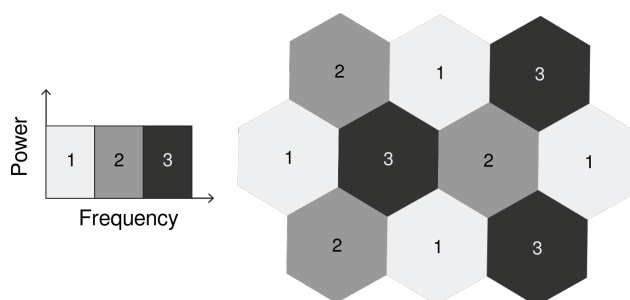


Figure 1. Scenario with reuse factor of 3. In this topology, no adjacent cells share bandwidth, i.e., their frequency bands are disjoint.

Fractional frequency reuse splits the cell into distinct regions with different reuse factors and transmission power. The goal is to improve the SINR, increasing the reuse factor without compromising the spectral efficiency. Users in the cell edge are usually the target of FFR-based techniques, since they are the most affected by co-channel interference.

The following section presents some FFR-based techniques, known in the 3GPP as ICIC techniques. These algorithms were selected because of the following reasons. They are standardized for LTE by 3GPP (but not limited to it, since they can be applied to other OFDMA-based systems, such as 5G NR), they are already implemented in the simulator used in this paper [19], and they represent a good range of ICIC strategies based on FFR that can be found in deployed networks [20]. The algorithms are fully described in [20–22].

3.1. ICIC Techniques

In order to compare the techniques, two algorithms without fractional reuse of the frequency are also considered: Full Frequency Reuse (NoOp) and Hard Frequency Reuse (HFR). The former has a reuse factor of 1, the latter has a reuse factor of 3, as illustrated in Figure 1, and both transmit with the same power for the whole bandwidth.

In scenarios with few users, the Full Frequency Reuse (NoOp) may present high data rates given the larger available bandwidth for each cell. However, higher interference is expected, and it may result in low SINR and high Packet Loss Ratio (PLR), especially for edge users. The HFR algorithm is more efficient in reducing interference, as all adjacent cells have disjoint bandwidths. However, each cell has only $1/3$ of the available bandwidth, which can severely reduce throughput, depending on the offered load [21].

3.1.1. Strict Frequency Reuse

The Strict Frequency Reuse (Strict FR) splits the bandwidth into two sub-bands, as illustrated in Figure 2. The cell-center User Equipments (UEs) are allocated to a common sub-band shared by all cells (reuse factor of 1) and the cell-edge UEs are allocated to a private sub-band (reuse factor of 3). Hence, cell-center users share bandwidth with neighboring cells, thus increasing spectral efficiency, but cell-edge users do not, thus reducing interference. Additionally, a higher power level is used for the private sub-band.

In order to determine whether a UE is allocated to the common or private sub-band, the ICIC algorithm uses a metric defined in 3GPP, the Reference Signal Received Quality (RSRQ). It indicates the quality of the signal and it takes into account various metrics, such as noise, power from interfering signals, and the number of allocated Resource Blocks (RBs). If the RSRQ reported by a user is higher than a threshold, the UE is allocated to the common sub-band (cell-center). Otherwise, it is allocated to the private sub-band (cell-edge). The RSRQ threshold is a user-defined parameter.

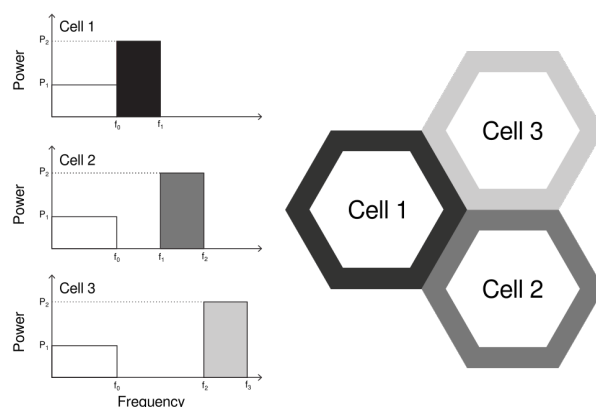


Figure 2. Bandwidth distribution and power allocation for the Strict frequency reuse algorithm in a cluster of three cells.

3.1.2. Soft Frequency Reuse

The Soft Frequency Reuse (SFR) also splits the bandwidth into two sub-bands. However, the sub-band allocated to cell-edge UEs is not private, as illustrated in Figure 3. However, cell-edge users only share bandwidth with cell-center users from adjacent cells. Consequently, the cell-edge sub-bands are disjoint. Thus, cell-edge UEs do not use the same frequency band as neighboring cells, and each occupies 1/3 of the available spectrum.

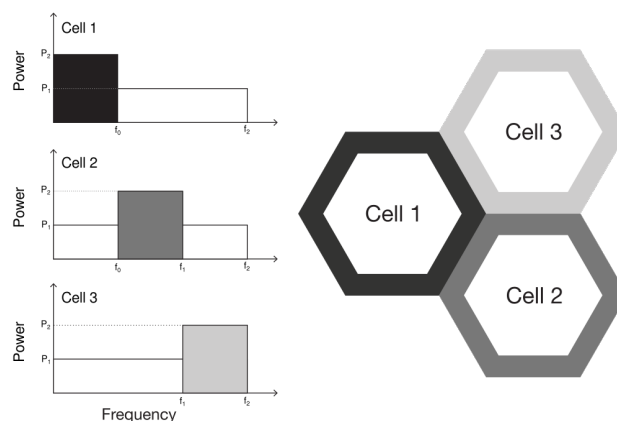


Figure 3. Bandwidth distribution and power allocation for the soft frequency reuse algorithm in a cluster of three cells.

The SFR can lead to better spectral efficiency when compared to the Strict FR, given that all cells can use the entire bandwidth. However, the SFR also increases the interference suffered by all users.

A higher power level is used for the edge sub-band, and a UE is considered in the cell-center if the reported RSRQ is greater than the threshold. Otherwise, it is allocated in the cell-edge sub-band.

3.1.3. Soft Fractional Frequency Reuse

The last technique is the Soft Fractional Frequency Reuse (SFFR). It divides the bandwidth into three distinct sub-bands: center, middle, and edge, as illustrated on Figure 4. The middle sub-band has a reuse factor of 1 and the edge sub-band has a reuse factor of 3. The center region reuses the frequency band from the edge of the adjacent cells.

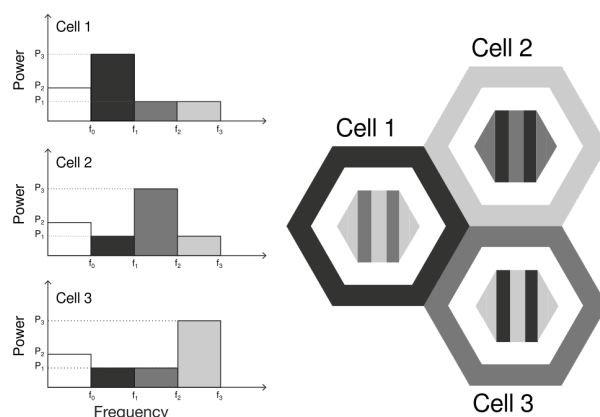


Figure 4. Bandwidth distribution and power allocation for the soft fractional frequency reuse algorithm in a cluster of three cells.

Each sub-band is served with a different power level, increasing from center to edge, and user allocation on sub-bands is also done based on the reported RSRQ.

4. ICIC in 3GPP Standards

LTE is a wireless communication standard introduced by 3GPP on Release 8. It has an all-IP network architecture and high flexibility on spectrum allocation, since OFDMA is the multiple access scheme on DL. Consequently, the eNB can allocate a UE to any sub-carrier in the frequency domain. This allocation is based on Resource Blocks (RBs) and each RB has 12 sub-carriers and 180 kHz of minimal bandwidth [23]. This flexibility is important to enable fractional reuse of the spectrum.

The X2 interface, introduced on LTE, is also important for ICIC. It enables signaling between eNBs, allowing for the ICIC algorithms to manage the RB and power allocation on neighboring cells [24].

Moreover, 3GPP's Release 8 also introduces implicit support to interference coordination [25], given that LTE has good control over time, frequency, and power resources. Release 9 presents studies that attempt to mitigate interference between macrocell and lower power nodes and, further on, Release 10 extends ICIC to the time domain as it introduces Enhanced ICIC (eICIC). Additionally, Release 11 presents the Further Enhanced ICIC (feICIC) in order to mitigate interference due to control signals.

Given this history, eICIC and feICIC are often considered as the evolution of ICIC. However, FFR-based techniques are still relevant based on recent studies, given that New Radio (NR), the new 3GPP standard for 5G systems, demands a flexible and efficient use of the spectrum [26]. For example, the authors of [27] propose a FFR-based ICIC scheme for NR. The algorithm dynamically allocates users suffering from interference to a set of Physical Resource Blocks (PRBs) that are not accessible to interfering cells.

5. Related Works

Metropolitan areas around the world have been growing significantly, alongside the number of users on wireless networks. This growth has encouraged the literature to widely discuss scenarios with a high density of users. Hence, the densification of the network is indicated as an efficient approach to better serve urban areas, including 5G networks [15]. Consequently, various works introduce ways of using small cells to boost the performance of scenarios with hotspots.

The authors of [28] propose a scenario where macrocells are populated with picocells centered on hotspots. The paper shows that the small coverage offered by the picocells causes an imbalance in user allocation, overloading the macrocells. Therefore, a dense ring of macrocell users is formed around the picocells. These users experience high interference coming from the picocells, and they are considered as victim users. They

propose two dynamic techniques of eICIC using two different metrics to improve the Quality of Service (QoS) of the victim users while muting picocell users.

A realistic dense scenario was evaluated in [29]. It simulated a wide urban area with 300 active users per km². Half of the users were uniformly distributed within the scenario, and the other half were uniformly distributed inside circular hotspots with a 40 m radius. For this scenario, the simulation results show that it is possible to achieve an average of 1 Gbps per user. However, they also noted that, given the year of publishing, the solution was not applicable due to its high cost and low energy efficiency.

The authors of [30] evaluated a scenario that deployed macrocells with three sectors each and four picocells per sector. Moreover, a fraction of the users were allocated in hotspots centered in the picocells. They proposed the use of time-domain ICIC, using Almost Blank Subframes (ABS) and Cell Range Extension (CRE) to improve performance and alleviate the macrocell load. The results show that the imbalance between cells is effectively countered with an offset value of 10 dB for the CRE. Besides, for each CRE offset, there was an optimum value for the ratio of protected subframes. Thus, if both parameters are properly configured within a certain range, the performance is almost the same.

Another technique used to improve the performance of hotspot scenarios is the use of multiple antennas. The authors of [8] proposed a HetNet scenario with small cells operating in the same frequency as macrocells. All users were allocated in hotspots and the small cells were located at the center of some of these hotspots, representing an intentional deployment of small cells. The solution focuses on mitigating the interference coming from the small cells through a beamforming scheme proposed in [7]. This scheme concentrates the transmission energy to the hotspots while creating transmission opportunities for users in other directions.

A scheduling algorithm that combines frequency allocation and beamforming (beams width and direction) is proposed in [31]. A homogeneous and a hotspot scenario were considered. The paper focused on maximizing throughput according to QoS requirements, and it compared the results to other approaches. The results show that the hotspot scenario is more challenging for all the algorithms. However, the proposed solution has better performance in terms of complexity and throughput.

The authors of [32] proposed a solution that served UEs in a hotspot scenario using virtual MIMO, i.e., user cooperation to enable spatial multiplexing. Compared to the traditional use of MIMO, this approach can avoid the costs of new antennas or access points dedicated to the hotspots. Furthermore, the signal processing was done in the mobile station, avoiding any new processing units. The proposed protocol presents better performance than traditional offloading techniques, but there is no discussion about privacy, the impact of the added signaling, or energy consumption on the user's side.

The literature has widely discussed solutions to the problems introduced by small cells, but, in general, only a few objectively show the negative effects related to the appearance of hotspots. In general, the deployment of small cells and the use of MIMO techniques demand investments in site-planning and infrastructure, and the latter usually requires channel estimation.

FFR is still relevant as an efficient approach to mitigate inter-cell interference in mobile networks [33–35]. Besides, recent studies still point out that FFR can be combined with different techniques to reduce interference, such as beamforming [34].

Hotspot scenarios are consistently studied for their importance regarding current and future mobile networks [36,37]. Still, no work has been found in the use of FFR as the primary solution to improve performance on hotspot scenarios without using small cells. FFR can be a good alternative for its simplicity and efficiency, even in heterogeneous scenarios with high user density [38]. This is one of the arguments that motivates the scientific hypothesis of this work.

6. System Model

6.1. Simulation Software

The network simulator 3 (ns-3) is the simulation tool used in this work [39]. It is open-source and, therefore, publicly available for development, education, and research activities. It is also modular, comprised of several models built primarily in C++ with some APIs available in Python. Its development is oriented by technical specifications from standard organizations, such as 3GPP and IEEE. Besides, it is well documented, with an active and collaborative community. For these reasons, it has been widely adopted for research and is the primary simulation tool for this paper.

6.2. LTE Module and ICIC on ns-3

We provide our proof-of-concept results using an LTE system model. The LTE module on ns-3 has two main components: the LTE model and the Evolved Packet Core (EPC) model. The former includes the E-UTRAN protocol stack, i.e., the RRC, PDCP, RLC, MAC, and PHY layers. These entities reside within the UE and the eNB nodes. The EPC model includes the core network functionalities, allowing end-to-end IP connectivity. All protocols and entities reside within the MME, S-GW, and P-GW nodes and partially within the eNB nodes. Figure 5 shows the LTE-EPC protocol stack for the data plane on ns-3 [40]. The only relevant simplification is the combination of the S-GW and P-GW functionalities into one node.

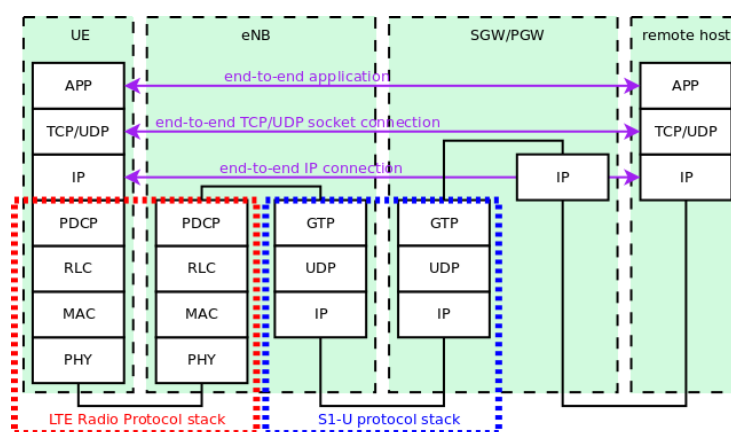


Figure 5. LTE-EPC data plane protocol stack (taken from the ns-3 documentation [40]).

There are currently seven ICIC algorithms implemented in the LTE module, described in [41]. The FFR algorithms act on scheduling, commanded by the MAC layer [40]. The algorithm is consulted and, depending on its rules for bandwidth allocation, it may allow (or not) the scheduling of a UE to a certain Resource Block Group (RBG). These algorithms have three main parameters that can be configured, as described below.

The power level of a sub-band is defined through a power offset between the Reference Signal (RS) and the Physical Downlink Shared Channel (PDSCH). Each sub-band has a variable that defines this offset in decibels. For example, the Strict FR has *CenterPowerOffset* and *EdgePowerOffset*.

Resource allocation in LTE is done through RBGs, but its size in number of RBs depends on the system bandwidth [25]. For example, for a system bandwidth of 100 RBs, each RBG has four RBs [25]. Hence, each sub-band has a variable that defines the number of available RBGs for DL and Uplink (UL), separately.

Section 3.1 introduced the RSRQ threshold, which determines user allocation on sub-bands. On LTE, RSRQ is measured in decibels and mapped into integer values before being reported [42]. On Release 8, these values range from 0 to 34. Hence, the variable *RsrqThreshold* only assumes values within this range.

6.3. Q-Learning on ns-3

In order to execute the proposed solution, a new class was added to ns-3. It implements the Q-Learning (QL) algorithm, which will be described in Section 10. Its first version was introduced by the authors of [43].

The simulation script includes both models described in Section 6.2. The EPC model enables the installation of applications in the UEs, which allows better control over the offered load and the appearance of hotspots during the simulation, considering the applications can be turned on and off at any given time.

All calculations related to the QL algorithm are made during *runtime*, while the simulation is executed. The algorithm also operates in the system during *runtime*, which guarantees that the network is able to adapt dynamically.

7. Preliminary Analysis A: Performance Comparison of ICIC Algorithms

The proposed solution of this paper is fully described in Section 10. However, before presenting the final results, we present three preliminary analyses that were important to guide some decisions regarding the Q-Learning algorithm. The first analysis is presented in this section, and it is a performance comparison between the ICIC techniques presented in Section 3.1. This study evaluates the impact of FFR on cell-edge UEs in a simple scenario. The algorithms were executed under the same conditions and their parameters assume the default values defined either on ns-3 or in the literature.

7.1. Evaluation Scenario

The scenario consists of 3 eNBs positioned in the vertices of an equilateral triangle with side equal to 1000 m, as illustrated in Figure 6, and each eNB is the center of a cell. This scenario was introduced by the authors of [19] to evaluate the ICIC techniques implemented on ns-3.

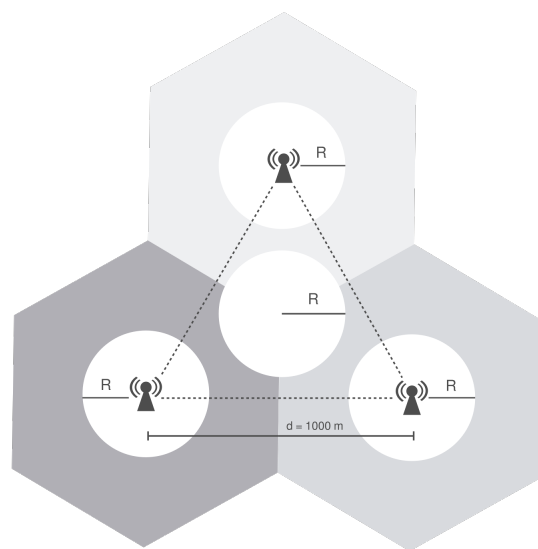


Figure 6. Scenario for preliminary analysis a: performance comparison of ICIC algorithms. The Radius R is set to 100 m in Scenario 1 and 500 m in Scenario 2, as described in Section 7.1.

Users are distributed uniformly inside circles controlled by a radius that can be configured as needed. There is a circle centered on each eNB and an additional circle in the triangle's centroid, which is the farthest location from each eNB.

User location within the circles is randomly selected from a uniform distribution, and all circles have the same radius. Given that a smaller radius leads to a more densely populated region (hotspot), it is possible to control user concentration. The number of users vary from 10 to 60 UEs on each circle and different circles always have the same

amount of users. Namely, if there are 30 users in the cell edge area, there are also 30 users on each cell-center, i.e., 120 in total. Therefore, the number of users range from 40 to 240.

Scheduling is made using the proportional fair algorithm and the link adaptation is based on SINR for a BER of $5 \cdot 10^{-5}$. The system bandwidth is 5 MHz, resulting in 25 RBs for each Transmission Time Interval (TTI). TTIs are the parameters related to the medium access in the radio link layer. At each TTI, the scheduling algorithm allocates RBs for all connected users. It is a parameter strongly related to access latency on LTE systems. Table 1 presents the configuration for each ICIC algorithm, suggested in [19]. Other simulation parameters are presented in Table 2.

The simulation was conducted in two scenarios:

- **Scenario 1:** UEs are concentrated in a 100 m radius, representing a classic hotspot scenario;
- **Scenario 2:** UEs are less concentrated, distributed over a 500 m radius circle.

Table 1. Parameters used for the comparison of the ICIC algorithms in preliminary analysis A [19].

Hard Frequency Reuse	
Bandwidth: cells 1 and 2	8 RBs
Bandwidth: cell 3	9 RBs
Strict Frequency Reuse	
Bandwidth (center/edge)	6 RBs each
RsrqThreshold	32
Power offset (center)	−6 dB
Power offset (edge)	3 dB
Soft Frequency Reuse	
Bandwidth (center)	25 RBs
Bandwidth (edge): cells 1 and 2	9 RBs
Bandwidth (edge): cell 3	9 RBs
RsrqThreshold	32
Power offset (center)	−6 dB
Power offset (edge)	3 dB
Fractional Soft Frequency Reuse	
Bandwidth (center/edge)	6 RBs each
RsrqThreshold (center)	37
RsrqThreshold (edge)	32
Power offset (center)	−6 dB
Power offset (middle)	−1.77 dB
Power offset (edge)	3 dB

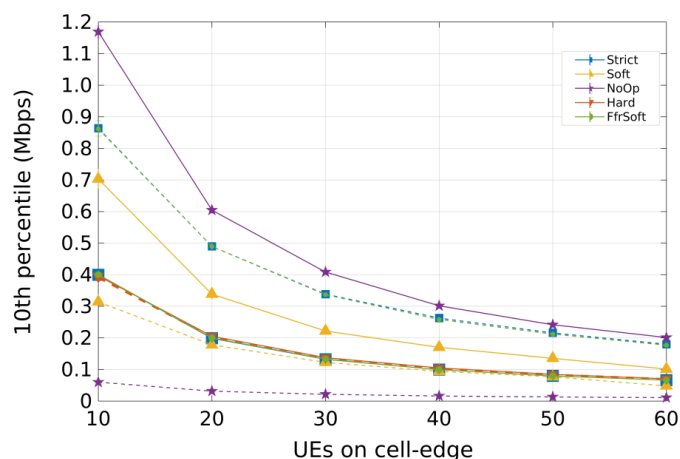
Table 2. Simulation parameters for preliminary analysis a: performance comparison of ICIC algorithms.

Parameter	Value
Bandwidth (RBs)	25
UE distribution	Uniform
Total UEs	40 to 240
Cell-edge UEs	10 to 60
Inter-eNBs distance	1000 m
Scheduling algorithm	<i>Proportional Fair</i>
Channel model	Friis model
Error model	MIESM
UE mobility	No mobility
Traffic model	Non-GBR TCP-based Video (Buffered Stream)

7.2. Results and Discussions

For Scenario 1 (concentrated users), Figures 7 and 8 present the 10th percentile throughput and the SINR Cumulative distribution function (CDF), respectively. The curves with dotted lines represent cell-edge users and the continuous represent cell-center users.

Regarding throughput, Figure 7 shows that the NoOp algorithm presents the best performance for cell-center users. All cells operate on the same frequency band; thus, more bandwidth is available, and a higher data rate is obtained. However, this is also the reason why NoOp has the worst performance for edge users, considering these UEs are more affected by interference.

**Figure 7.** Scenario 1 from preliminary analysis A: 10th percentile throughput for concentrated users ($R = 100\text{m}$). Each curve represents a different ICIC algorithm.

Strict FR and SFFR have similar performance. They present the best results for cell-edge users, but limited performance for cell-center users. These algorithms allocate the same number of RBs for each region. Hence, cell-center users only have 25% of the bandwidth available (Table 1). The HFR algorithm has similar performance for both center and edge users, since it does not divide the cell into distinct regions. Users do not suffer from inter-cell interference, but the smaller bandwidth leads to lower throughput and both Strict FR and SFFR have better performance for all UEs.

The SFR obtains good performance for cell-center users but poor performance for cell-edge users. The high number of RBs reserved for cell-edge users was not enough to balance Inter-Cell Interference (ICI), since the cell-edge bandwidth is shared with cell-center regions from adjacent cells.

Figure 8 indicates that cell-edge users have lower SINR for all algorithms, and the worst results for both sets of users are from SFR and NoOp. HFR has the best performance for cell-center users, but Strict FR and SFFR have better performance for cell-edge users while also maintaining good performance on cell-center, similar to HFR. Moreover, considering the lower spectral efficiency, as discussed before, the HFR algorithm is not able to provide good throughput levels.

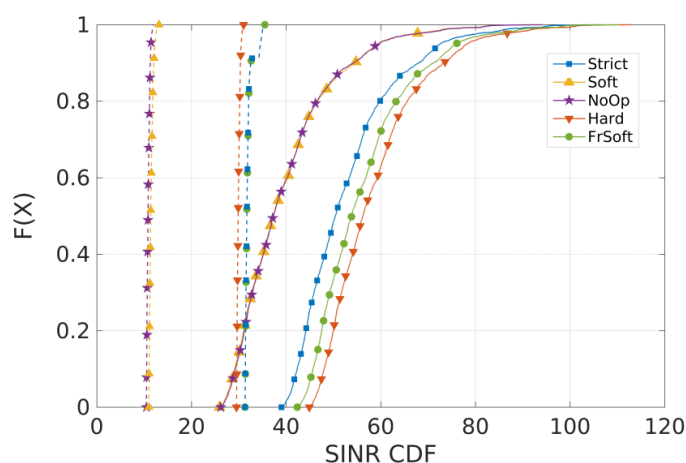


Figure 8. Scenario 1 from preliminary analysis A: SINR CDF for concentrated users ($R = 100$ m). Each curve represents a different ICIC algorithm.

Figures 9 and 10 show the results for Scenario 2, with less concentrated users. Concerning throughput (Figure 9), there is a significant performance loss for NoOp and SFR, mainly because cell-center UEs can interfere with cell-edge UEs. As a result, spreading the users can increase interference, due to higher occupation of the cell-center region. Additionally, Strict FR and SFFR still have the best throughput levels for cell-edge users.

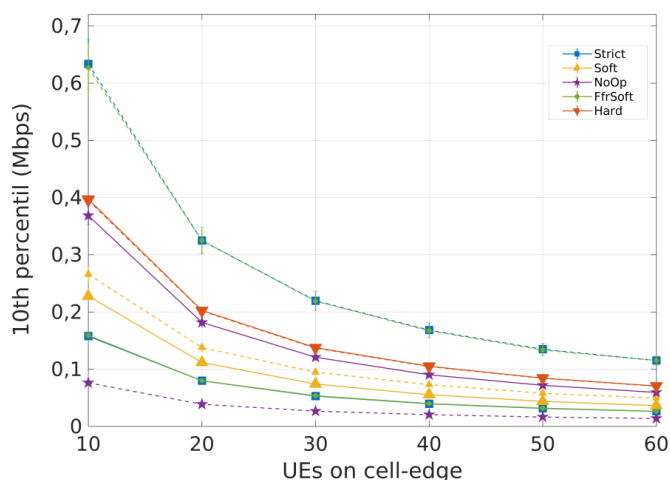


Figure 9. Scenario 2 from preliminary analysis A: 10th percentile throughput for less concentrated users ($R = 500$ m). Each curve represents a different ICIC algorithm.

Regarding SINR results, Figure 10 shows that NoOp and SFR still have the worst performance in all cases and the Strict FR and SFFR algorithms maintain good performance

on cell-edge. Nevertheless, all algorithms present lower SINR on cell-edge, since, in this case, spreading the users led to more populated cell-edge regions.

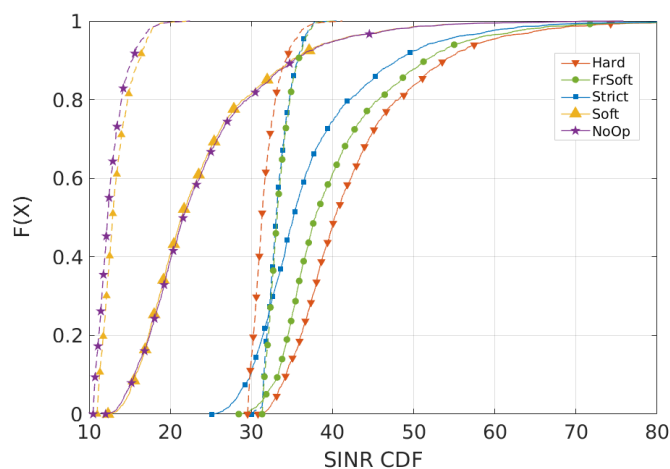


Figure 10. Scenario 2 from preliminary analysis A: SINR CDF for less concentrated users ($R = 500$ m). Each curve represents a different ICIC algorithm.

Some important conclusions can be outlined:

- FFR-based ICIC techniques can improve the performance of a mobile network, given the high interference suffered by its users. It is possible to enhance cell-edge performance without compromising cell-center users;
- Simply changing the user density can strongly impact on system performance, especially if the bandwidth is divided into sub-bands. The concentration or spreading of the users can result in different occupation of the sub-bands, leading to overloaded or nearly empty sub-bands;
- There is no scheme that performs best in every situation. However, the Strict FR and the SFFR schemes have a good performance in different scenarios, and they are both efficient at reducing ICI, especially for cell-edge users;
- The Strict Frequency Reuse (Strict FR) algorithm has the best compromise between performance and complexity, considering SFFR has a higher number of sub-bands, which leads to more complexity when adjusting the parameters. Therefore, the Strict FR is the object of the study presented in the following sections.

8. Preliminary Analysis B: Factorial Design using the Strict Frequency Reuse Algorithm

The second step towards the proposed solution consists of selecting one of the ICIC algorithms to evaluate the impact of its most relevant parameters on system performance. This analysis has two further steps: a 2k factorial design and a full factorial design.

These studies will be conducted in the Strict FR algorithm for its good compromise between complexity and performance.

8.1. Evaluation Scenario

The proposed scenario is very similar to the one described in Section 7.1, as illustrated in Figure 11. It consists of three eNBs positioned in the vertices of an equilateral triangle with side equal to 1000 m. A total of 80 users are randomly positioned in the entire scenario according to a uniform distribution. They do not have mobility and each UE is served by the closest eNB. The simulation parameters are presented in Table 3.

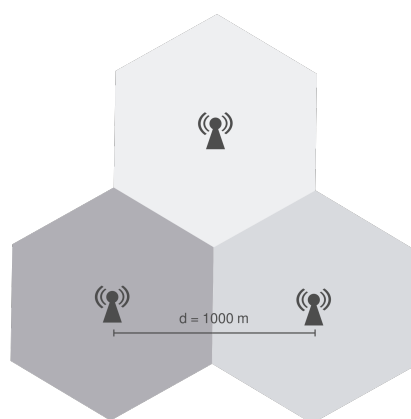


Figure 11. Scenario for preliminary analysis B, which is similar to the scenario in Figure 6, but without the radius R , since all users are randomly distributed according to a uniform distribution.

Table 3. Simulation parameters for preliminary analysis B: parametric analysis of strict frequency reuse.

Parameter	Value
Bandwidth (RBs)	25
UE distribution	Uniform
Number of UEs	80
Distance between eNBs (m)	1000
Scheduling algorithm	<i>Proportional Fair</i>
Channel model	Friss model
Error model	MIESM
UE mobility	No mobility
Traffic model	Non-GBR TCP-based Video (Buffered Stream)

8.2. 2^k Factorial Design

The 2^k factorial design (2^k factor) is a specific case of factorial design that identifies which parameters of an experiment have a significant effect on the desired output [44]. It provides a better understanding of the system, and can simplify future analysis. For example, assume that the 2^k factor is conducted for an experiment with a parameter called P . If the 2^k factor determines that P is not relevant, a following full factorial design will need fewer repetitions, since P does not need to change. A Full Factorial Design consists of replicating an experiment for all possible combinations of parameters.

To perform the 2^k factor, k factors (parameters) are chosen, and they assume only two distinct values, preferably a low and a high value close to the upper and lower limits of the parameter's range. Besides, it is usually followed by an Analysis of Variance (ANOVA) that validates the results by statistically rejecting or not the null hypothesis, which is a given parameter does not affect the desired output. ANOVA tests this hypothesis by comparing variances.

Consider the value F_0 in Equation (1). This value is calculated for each parameter, and it is defined as the division between two Sum of Squares (SSs), which is a non-biased estimator of a population's variance [44]. For example, if a simulation that has a parameter A is repeated n times for each value of the parameter, SS_{treat} is the SS that measures the variation between the different values of parameter A and SS_{error} is the SS that measures the variation between the n results of a single value of A . If the change of a given parameter

has a relevant impact on the desired output, the null hypothesis is rejected, and SS_{treat} is bigger than SS_{error} [44]. Therefore, the bigger F_0 is, the bigger is the parameter's impact.

$$F_0 = \frac{SS_{treat}}{SS_{error}} \quad (1)$$

However, it is recommended to define a threshold for F_0 that guarantees the parameter's relevance. Since SS_{treat} and SS_{error} are both chi-square random variables by definition [45], F_0 has a Fisher–Snedecor distribution [44,46]. Defining a level of significance $\alpha = 0.001$, for the evaluated parameters, the threshold value for F_0 is 10.83. Thus, F_0 is calculated for each parameter. If F_0 is greater than 10.83, the parameter has a significant impact on the output.

Therefore, the 2^k factorial was conducted for the Strict FR's parameters: *CenterPowerOffset*, *RsrqThreshold*, and *BandwidthDistribution*. As detailed in Section 6, they define the power level of the common sub-band, the threshold for allocating users on sub-bands and the number of RBs on each sub-band, respectively. In order to investigate all users and users with the worst performance (most likely cell-edge users), the targeted output is the average throughput and its 10th percentile.

Tables 4 and 5 present the calculated F_0 for each parameter and for the interaction between parameters. The nomenclature $X * Y$ represents the interaction between parameters X and Y . *CenterPowerOffset* is the only parameter that does not have a significant impact on the average or 10th percentile throughput. This is due to how bandwidth is allocated in the Strict FR. The cell center of each cell does not share bandwidth with the cell-edge of any cell. Hence, users from different cells allocated to the same bandwidth might not be close enough to cause significant ICI.

Table 4. F_0 results for average throughput.

Parameter (Average Throughput)	F_0
CenterPowerOffset (A)	0.084
RsrqThreshold (B)	39040.4
BandwidthDistribution (C)	4301.4
BandwidthDistribution * CenterPowerOffset	0.0015
BandwidthDistribution * RsrqThreshold	6264.7
RsrqThreshold * CenterPowerOffset	0.0012
A * B * C	0.00032

Table 5. F_0 results for the 10th percentile.

Parameter (10th Percentile)	F_0
CenterPowerOffset (A)	0.00059
RsrqThreshold (B)	2688.37
BandwidthDistribution (C)	6359.40
BandwidthDistribution * CenterPowerOffset	0.035
BandwidthDistribution * RsrqThreshold	695.91
RsrqThreshold * CenterPowerOffset	0.0033
A * B * C	0.00034

8.3. Full Factorial Design

After identifying which parameters are relevant, a full factorial design repeats the simulation to vary these parameters. The purpose of this investigation is to evaluate how

system performance is affected by the Strict FR's parameters, in order to identify the best configurations. Preliminary tests indicated that 100 repetitions are enough for the proposed scenario to provide a good confidence interval. Moreover, given that CenterPowerOffset represents a power level and does not impact on throughput, it is fixed on its minimum possible value, lowering energy consumption. Each evaluated parameter is detailed in Table 6.

Table 6. Full factorial design: parameter configurations.

Parameter	Values
RsrqThreshold	24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34
Common/Private Sub-band (RBs)	Bandwidth Distribution 1: 6 / 18 Bandwidth Distribution 2: 12 / 12 Bandwidth Distribution 3: 18 / 6

Figures 12 and 13 present the results for the average and 10th percentile throughput, respectively. Both have a 95% confidence interval. Lower values of RsrqThreshold result in better performance if more bandwidth is allocated to cell-center users. Consequently, bandwidth distribution 1, which allocates 6 RBs for cell-center and 18 RBs for cell-edge users (6 RBs for each cell edge), has the worst performance. This is due to more users being allocated to the cell center as RsrqThreshold gets lower, causing the common sub-band to overload.

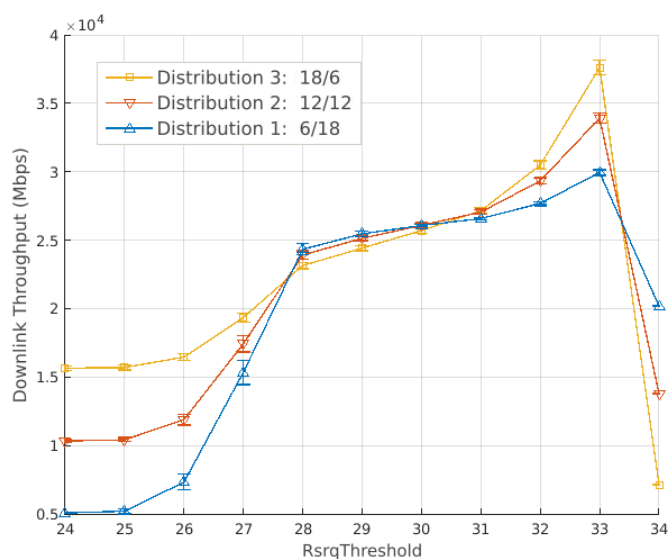


Figure 12. Full factorial design using the strict frequency reuse: results for the average throughput. Each curve is a different bandwidth distribution.

Moreover, increasing the RsrqThreshold yields higher throughput for all distributions (Figure 12) until each curve reaches its peak. At this point, the distribution with more band allocated to the center still has the best performance. From that value, the private sub-band gets overloaded, which results in performance loss. This loss is more severe for distributions with less bandwidth allocated to the edge.

Figure 13 has similar behavior to Figure 12, since increasing the RsrqThreshold results in higher throughput. However, each curve starts decreasing at different values. Distributions with less bandwidth to the edge sub-band have their peak at lower RsrqThreshold values. At RsrqThreshold = 33, Figure 12 indicates that throughput is better for distributions with more bandwidth allocated to the center. Figure 13 indicates the opposite, and bandwidth distributions 1 and 2 have similar performance on RsrqThreshold = 32.

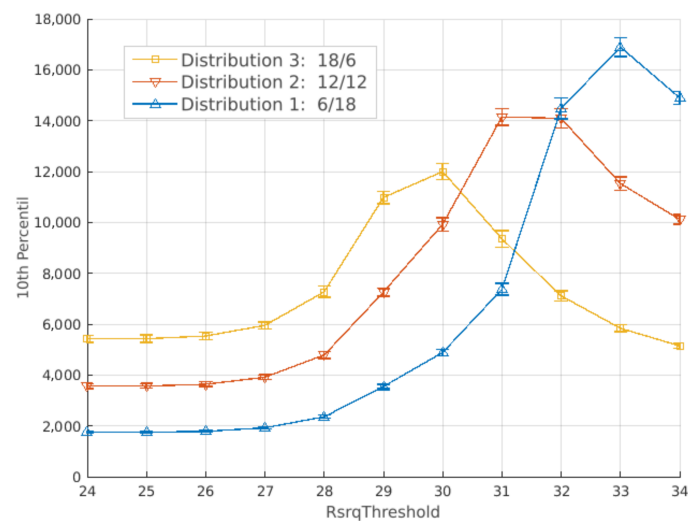


Figure 13. Full factorial design using the strict frequency reuse: results for the 10th percentile throughput. Each curve is a different bandwidth distribution.

Lastly, Figure 14 presents the SINR CDF for $RsrqThreshold = 3$. At this value, more bandwidth allocated for cell-edge users result in better SINR.

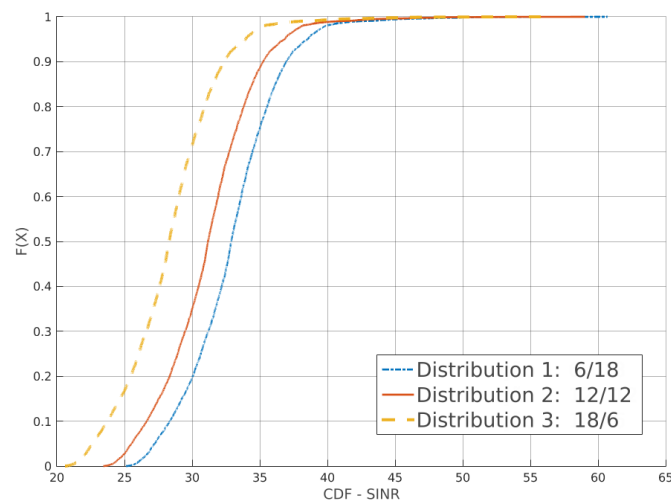


Figure 14. Full factorial design using the strict frequency reuse: results for the SINR CDF. Each curve is a different bandwidth distribution.

These results show a clear compromise between the average user performance and the performance of cell-edge users. Besides, in a scenario that hotspots appear unpredictably, certain regions can also become overloaded. Since the $RsrqThreshold$ has direct impact in user allocation, its dynamic variation can efficiently mitigate the performance loss caused by the hotspots.

9. Preliminary Analysis C: Evaluating the Hotspot Scenario

The last set of investigations before presenting the proposed solution evaluates the impact of hotspots on system performance. The goal is to compare a scenario with hotspots to a scenario with users distributed uniformly.

9.1. Evaluation Scenario

The evaluation scenarios are similar to those in Sections 7.1 and 8.3, since there are three eNBs equally distanced by 1000 meters. In this analysis, scenarios 1 and 2 have uniformly distributed users, while scenario 3 also has 5 hotspots with 15 users each. Figure 15 illustrates the approximate position of each hotspot, which are not random. The differences between these scenarios are summarized in Table 7.

There are two scenarios without hotspots, but with different amounts of users. As a result, it is possible to evaluate the impact of hotspots that increase (or not) the total amount of users in the system. Other simulation parameters are presented in Table 8.

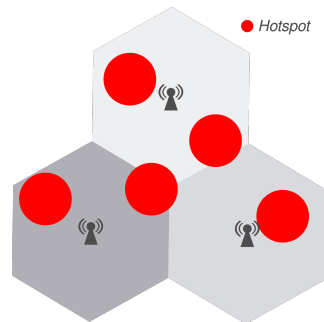


Figure 15. Scenario for preliminary analysis C with the approximate position for each hotspot, which are not randomly located.

Table 7. Difference between hotspot scenarios.

Scenario 1
60 uniformly distributed users and no hotspots
Scenario 2
135 uniformly distributed users and no hotspots
Scenario 3
135 users: 60 uniformly distributed users and 5 hotspots with 15 users each

Table 8. Simulation parameters: hotspot scenarios.

Parameter	Value
Bandwidth (RBs)	100
UE distribution	Uniform / Hotspots
Number of UEs	60 / 135
Distance between eNBs (m)	1000
Scheduling algorithm	<i>Proportional Fair</i>
Simulation duration	6000 subframes
Channel model	Friis Model
Error model	MIESM
UE mobility	No mobility
Traffic model	Non-GBR TCP-based Video (Buffered Stream)

9.2. Simulation Results

Figure 16 presents the simulation results in terms of the 10th percentile of throughput. The dashed lines represent the scenario with hotspots (Scenario 3). The simulations were executed using the Strict FR algorithm, with the variation of its relevant parameters, according to Section 8 (*RsrqThreshold* and *BandwidthDistribution*). The former ranges between 30 and 33, and the latter is evaluated in two configurations. The first allocates 52 RBs to the common sub-band and 16 RBs to the private sub-band. The second allocates 28 RBs to the common sub-band and 24 RBs to the private sub-band.

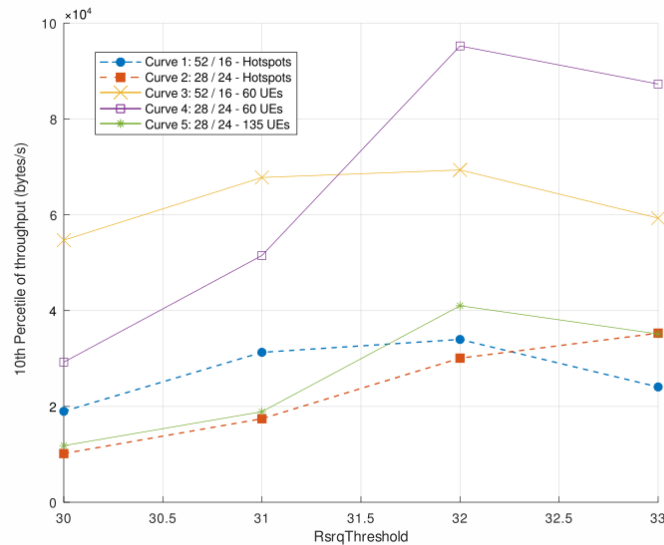


Figure 16. Simulation results in terms of the 10th percentile of throughput. The dashed lines represent the scenario with hotspots.

The 10th percentile of throughput (Figure 16) indicates a significant performance loss due to the hotspots. The curves with the same distribution (curves 2, 4, and 5) show that the scenario with hotspots has lower throughput, especially if compared to curve 4, which has fewer users.

Even though the performance loss is expected, given the increase in the number of users, both cases discussed above are important. Section 2 discussed scenarios naturally characterized by hotspots. In most of these scenarios, the hotspots also increase the total number of active users in the system, such as football games or music festivals. However, even if the change is only in user concentration, there is a relevant performance loss, as presented in Figure 16, given that curve 2 has the worst performance than curve 5.

10. Proposed Solution: Background, Implementation, and Simulation Results

The results presented in Section 7 show that FFR-based techniques efficiently mitigate ICI, improving throughput and SINR. It also shows that no algorithm performs best in every situation, but the Strict FR has a good compromise between performance and complexity. Moreover, not all of its parameters have a significant impact on throughput and SINR, according to Section 8.

Results from Section 9 indicate that hotspots can have a significant impact on system performance, and different states of the system may require different configurations of the Strict FR parameters to achieve improved performance. Furthermore, Section 2 discusses challenges related to hotspot scenarios and highly dynamic urban areas. In such scenarios, techniques that allow the system to adapt dynamically may increase performance.

Therefore, this section presents our proposed solution for scenarios where hotspots appear unexpectedly. This solution applies Machine Learning (ML) techniques to dynamically regulate the *RsrqThreshold* of the Strict FR.

10.1. Machine Learning Techniques

Machine Learning (ML) is a subset of artificial intelligence that builds mathematical models to find patterns in data sets. The goal is to make decisions without human assistance. They have been widely discussed in the literature for their ability to adapt to changing scenarios and reduce site planning. For instance, the authors of [47] present various ML algorithms and how they have been employed to coordinate co-channel interference.

Moreover, these techniques can be classified as supervised, non-supervised, and Reinforcement Learning (RL). Supervised learning algorithms are akin to a supervisor that holds the desired knowledge. These algorithms adapt based on inputs previously known by the supervisor (labeled data sets). Non-supervised algorithms do not have labeled inputs. Hence, they aim to find hidden patterns or similarities within the data. On the other hand, RL continuously interacts with the system, and each decision produces a reward or punishment that serves to update the algorithm parameters. The goal is to maximize or minimize the reward or punishment [48].

For the proposed scenario, RL is the most appropriate paradigm. Supervised learning requires prior knowledge and dedicated time for training, which is not desirable for a system that demands constant re-adapting. Non-supervised learning does not require prior knowledge, but RL is a better option for this task, considering the algorithm learns from each action taken using immediate rewards. RL is an efficient paradigm for interaction with uncertain environments [49], enabling real-time decision making.

10.2. Reinforcement Learning

Reinforcement Learning (RL) is a ML paradigm that has a learning agent that attempts to reach an objective through trial and error, i.e., through constant interactions with the system. The agent is not instructed on which action should be taken, but it must be capable of observing the state of the system and take actions that can change it.

As a consequence of the action, the system provides a numerical reward and informs its new current state. The agent must identify which actions result in bigger rewards in any state of the system. Hence, the algorithm estimates the reward for each action a_n for each state s_n . However, the estimated value is not based on the current action alone. It considers all of the actions taken so far [49].

Among RL approaches, Q-Learning (QL) has been constantly discussed in the literature, especially because it does not require estimating the dynamics of the environment (it is model-free) [47].

For example, the authors of [50] attempt to maximize system performance in an ultra-dense heterogeneous scenario. They propose a dynamic resource allocation scheme using centralized and distributed QL. The authors of [51] use QL in a HetNet scenario. Each cell learns optimum values for the Cell Range Bias (CRB) and the DL power transmission level. The results show a 125% improvement compared to static ICIC techniques. Additionally, the authors of [52] evaluate a scenario where a stadium is served by 78 eNBs. They propose an improved QL algorithm that decreases the convergence time.

Our previous works include [43,53]. They present solutions that use ABS and QL to coordinate the shared transmission between LTE and Wi-Fi in the 5 GHz bandwidth. Therefore, this paper applies the methodology previously tested, especially regarding the prototyping of a dynamic solution of Radio Resource Management (RRM) using QL.

10.3. Q-Learning

Q-Learning (QL) [54] operates by creating a table of q-values using a $Q(s, a)$ function that estimates the rewards for each state/action pair. $Q(s, a)$ is the q-value of action a when the system state is s . The q-value indicates the desirability of taking action a when in state s . The higher the q-value, the more desirable it is.

The values estimated by this function are updated with each iteration, according to Equation (2) [49]:

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha [R_{t+1} + \gamma \max_a Q(S_{t+1}, a) - Q(S_t, A_t)], \quad (2)$$

where R_{t+1} is the reward from action A_t when the state is S_t ; α is the learning rate, which determines the impact of new information in the update of the q-values; and γ is a discount factor, which determines the impact of future rewards. Its value ranges from 0 to 1, and values close to zero indicate that future rewards are not important. In this case, immediate rewards have a greater impact in the learning process.

Therefore, the algorithm's goal is to find the optimum $Q(s, a)$ table, which gives an optimum policy π^* . This policy corresponds to choosing action a , given a state s , as defined in Equation (3):

$$\pi^*(s) = \arg \max_a Q(s, a). \quad (3)$$

The author of [55] demonstrates that, given enough samples, the QL algorithm converges to a function directly derived from the Bellman equations, which is the classic solution of a Markov Decision Process (MDP). The key to that convergence is a Markovian Process named Action-Replay Process (ARP). It is possible to use the weak law of large numbers to prove that the policy π gets closer to the optimum policy π^* for a growing number of samples.

10.4. The Q-Learning Implementation

As a consequence of the discussions and exploratory analysis presented, the QL-based solution aims at maximizing the average SINR to mitigate the negative impact of ICI on a dynamic scenario. The algorithm controls the *RsrqThreshold*. According to the results shown in Section 8, the *RsrqThreshold* is the parameter that has the biggest impact on average throughput. Hence, it was the first option for the QL algorithm. However, future works include the control of both the *RsrqThreshold* and the *BandwidthDistribution*, which is also an important parameter for the ICIC performance parameter, which dictates user allocation in the sub-bands defined by the Strict FR.

The QL algorithm operates as a centralized solution, namely that every decision affects all cells in a cluster. The ICIC techniques presented in Section 3.1 split the bandwidth for clusters formed by three cells. Therefore, each change in the *RsrqThreshold* affects all three cells of the cluster. If each cell had an independent QL algorithm, they would be competing, as each action would affect the whole cluster, based only on the data from one cell.

The QL algorithm needs the following parameters:

- A set of available actions, $A = a_1, a_2, \dots, a_n$;
- A set of possible states of the system, $S = s_1, s_2, \dots, s_n$;
- A $Q(s, a)$ matrix to store the estimated rewards;
- α and γ .

The algorithm was configured with six states and six actions. The metric that defines the state is the SINR; however, this metric may vary. For example, depending on the service targeted, it could be throughput, PLR, delay, or a combination of those, as presented in Equation (4) [56]:

$$M_a = \sum_{i=1}^n w_i f_i(M_i), \quad (4)$$

where w_1, \dots, w_n are the weights of each metric (M_1, \dots, M_n) and f_i is a normalization function that maps the metrics to values between 0 and 1.

The SINR was chosen because it directly indicates the impact of interference in users. Hence, the average SINR defines the states of the system and it is used as the reward from each state/action pair. Through experimental analysis to characterize the scenarios, the states are defined as follows:

- State 1: $SINR < 24 \text{ dB}$;

- State 2: $24 \text{ dB} < \text{SINR} \leq 27 \text{ dB}$;
- State 3: $27 \text{ dB} < \text{SINR} \leq 29 \text{ dB}$;
- State 4: $29 \text{ dB} < \text{SINR} \leq 30 \text{ dB}$;
- State 5: $30 \text{ dB} < \text{SINR} \leq 32 \text{ dB}$;
- State 6: $\text{SINR} > 32 \text{ dB}$.

Additionally, the set of available actions corresponds to RsrqThreshold values: $A = \{29, 30, 31, 32, 33, 34\}$. These values were selected based on the results presented in Section 8, given that this parameter has a bigger impact on system performance when operating on that range. The matrix $Q(s, a)$ initially has all values equal to zero and the values for α and γ are 0.4 and 0.5, respectively. This choice is based on exploratory simulations and related works [43] and the $Q(s, a)$ values are calculated according to Equation (5):

$$Q(s, a) \leftarrow (1 - \alpha)Q(s, a) + \alpha \left[r + \gamma \max_a Q(s', a) \right] \quad (5)$$

Algorithm 1 presents a pseudo-code of the proposed solution. The QL algorithm is executed each 40 ms. This interval is inspired by the study presented in [43]; however, exploratory simulations were conducted to assure that the SINR samples collected during this interval are sufficient to represent the average SINR of the system.

As mentioned previously, the QL seeks an optimum solution through a succession of actions. Thus, at any given time, the algorithm may get stuck on a set of state/action pairs without exploring other options (similar to an optimization algorithm that gets stuck on a local minimum). To address this issue, the parameter ϵ may be used to control the algorithm's level of exploration, making the agent choose a random action eventually.

10.5. Evaluation Scenario

The scenarios are similar to that of Section 9. There are three eNBs equally distanced by 1000 m, and the users are allocated in two manners. The first group of users is randomly allocated in the entire scenario, using a uniform distribution. The second group is allocated on hotspots with a 100 m radius, and there are five hotspots with fixed locations, as illustrated in Figure 15. This positioning aims at modeling different situations: close and far from a eNB and between multiple eNBs.

Algorithm 1: Pseudo-code of the proposed *Q-learning* algorithm.

```

1 Initialize
2 for  $s \in S, a \in A$  do
3   | Initialize Q-table with all  $Q(s, a)$  equal to 0.
4 end
5 Estimate initial state  $s$ .
6 Learning:
7 Loop
8 Generate random number  $r \in U(0, 1)$ 
9 if  $r < \epsilon$  then
10  | Randomly choose action  $a \in A$ ;
11 else
12  | Choose action  $a \in A$  according to Equation 3;
13 end
14 Execute action  $a$ ;
15 Receive immediate reward  $r, \text{SINR}_{\text{média}}$ ;
16 Observe next state  $s' \in S$ ;
17 Update Q-table according to Equation 5;
18  $s = s'$ 
19 end loop

```

As described in Section 2, this paper uses the term hotspot to describe regions with high user density. It does not mean a new AP. Hence, hotspot users do not have a dedicated AP. All users are served by the closest eNB.

Table 9 presents the difference between the scenarios used to evaluate the QL algorithm. **Scenario 1** has 60 users distributed uniformly and 10 users on each hotspot. A total of 52 RBs are allocated for the common sub-band, and 16 RBs are allocated to the private sub-band. **Scenario 2** has the double of users on each hotspot and 2/3 of the uniform-distributed users when compared to Scenario 1. Hence, there are 40 users distributed uniformly and 20 users on each hotspot, which results in a higher number of users in the system. A total of 28 RBs are allocated for the common sub-band, and 24 RBs are allocated to the private sub-band. Therefore, the available bandwidth for users allocated in the common sub-band is significantly smaller. Other simulation parameters are summarized in Table 10.

Table 9. Evaluations scenarios for proposed solution.

Scenario 1
60 users uniformly distributed
10 users on each hotspot
BandwidthDistribution: 52/16
Scenario 2
40 users uniformly distributed
20 users on each hotspot
BandwidthDistribution: 28/24

Table 10. Simulation parameters for the QL algorithm evaluation.

Parameter	Value
Bandwidth (RBs)	100
UE distribution	Uniform / Hotspots
Number of UEs	110 / 140
Distance between eNBs (m)	1000
Scheduling algorithm	<i>Proportional Fair</i>
Simulation duration	60,000 subframes
Channel model	Friis model
Error model	MIESM
UE mobility	No mobility
Traffic model	Non-GBR TCP-based Video (Buffered Stream)

A hotspot (HS) begins to transmit and receive data every 10 s to evaluate the algorithm's ability to adapt. From 0 to 10 s, only the uniformly distributed users are active. Between 10 and 20 s, one of the hotspots is also transmitting/receiving data. From 30 s, another hotspot becomes active, and so forth. The change on the scenario during simulation is presented in Table 11. Before discussing the results, the following section details how data are collected on the ns-3 simulator to enable the QL calculations.

Table 11. Active users at each simulation interval.

Interval	Scenario 1 (UEs)	Scenario 2 (UEs)
0 to 10 s	60 + 0 on HS	40 + 0 on HS
10 to 20 s	60 + 10 on HS	40 + 20 on HS
20 to 30 s	60 + 20 on HS	40 + 40 on HS
30 to 40 s	60 + 30 on HS	40 + 60 on HS
40 to 50 s	60 + 40 on HS	40 + 80 on HS
50 to 60 s	60 + 50 on HS	40 + 100 on HS

10.6. Data Collection on ns-3

Researchers using ns-3 can use a mechanism called *tracing* to capture any data generated by simulations. With this subsystem, the user can connect a function to any *trace source*, which outputs the data to the connected function when a condition is satisfied (e.g., every time a user transmits). This paper uses the trace source *ReportCurrentCellRsrpSinr*. It returns linear SINR values for all users from a specific eNB at each LTE subframe (every 1 ms). This trace source is implemented in the PHY layer from the LTE module. For the throughput, the trace source *RxPDU* notifies the amount of received Protocol Data Units (PDUs) by the Radio Link Control (RLC) entity.

The function connected to the first trace source stores SINR values from all eNBs and, every 40 ms, calculates the average SINR of the system on that interval (accumulated sum divided by the number of samples). Similarly, the second trace source is connected to a function that stores the data in text files. These files contain the number of PDUs and bytes transmitted and received. The throughput is calculated by dividing the total amount of received bits by the transmission interval in seconds.

10.7. Proof-of-Concept Simulation Results

The results are presented using SINR and throughput. Moreover, users are split in three groups: all users, users in hotspots, and users outside hotspots. Figures 17 and 18 present the SINR percentage gains for all groups of users. Figures 19 and 20 present throughput results in Mbps and Figures 21 and 22 present the throughput percentage gain.

Given the wide range that SINR is reported in linear scale, it is more practical to calculate the gain on a logarithmic scale. Hence, it is executed using dB values, as follows:

$$Gain = \left(\frac{SINR_{QL}}{SINR_{noQL}} - 1 \right) \times 100\% \quad (6)$$

where $SINR_{QL}$ is the system average SINR in dB when our solution is applied, and $SINR_{noQL}$ is the system average SINR without it. Thus, the *Gain* measures how our proposal improves the SINR in the dB scale.

According to Figure 17, all intervals and groups of users present relevant gain, even when there are no active hotspots (0 to 10 s). Additionally, as the number of active hotspots increases, the SINR degrades more severely, and the gain obtained tends to grow, since QL operates to improve the SINR. Hence, the two intervals with more active hotspots present over 80% gain. In addition, for most intervals, the gain is greater for hotspot users, which is interesting, considering the algorithm was not configured to act based on the performance of these users.

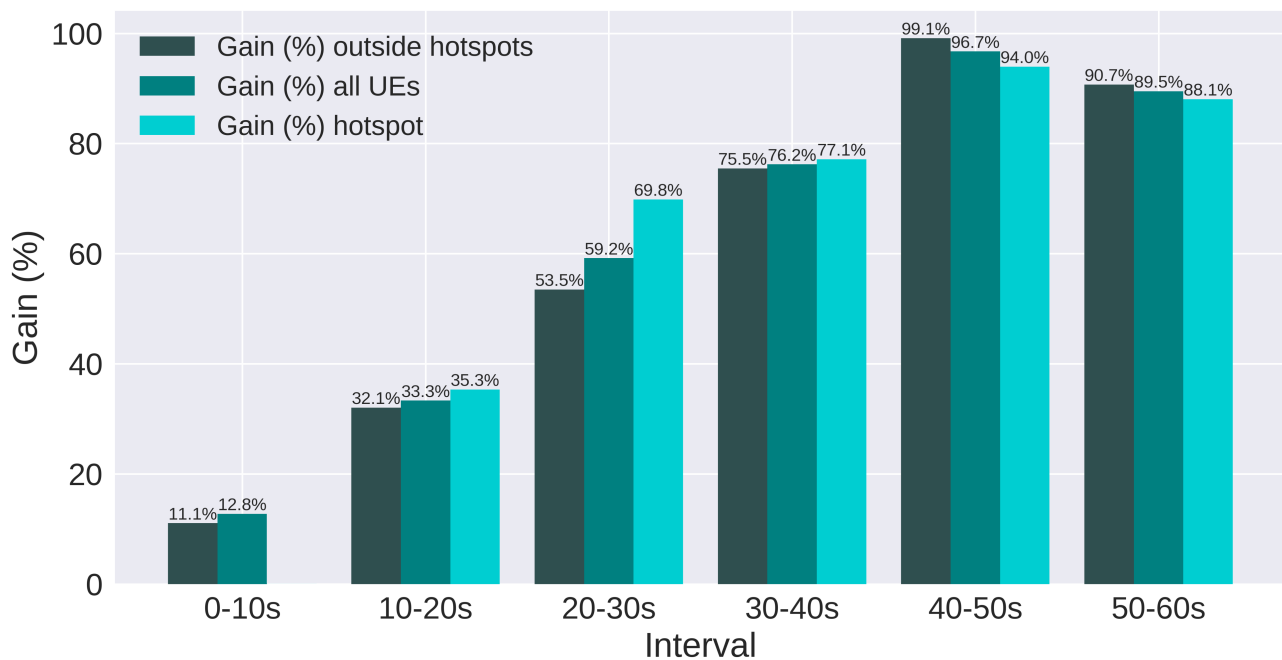


Figure 17. Scenario 1 of the proposed solution: SINR Gain (%) for each group of users.

Figure 19 presents throughput results for the same scenario. When the QL is not active, RsrqThreshold is fixed on 32. In this plot, three important results are combined: the throughput of users outside the hotspots (bottom bars), the throughput of hotspot users (top bars), and the average throughput of all users (combination of both bars). Each group of two parallel bars corresponds to a 10 s interval during the simulation (Table 11), and the bars on the left present the results when the QL is not active. The plot shows that the SINR gain resulted in throughput gain for all groups of users. However, the gain for users outside the hotspots is smaller when compared to hotspot users.

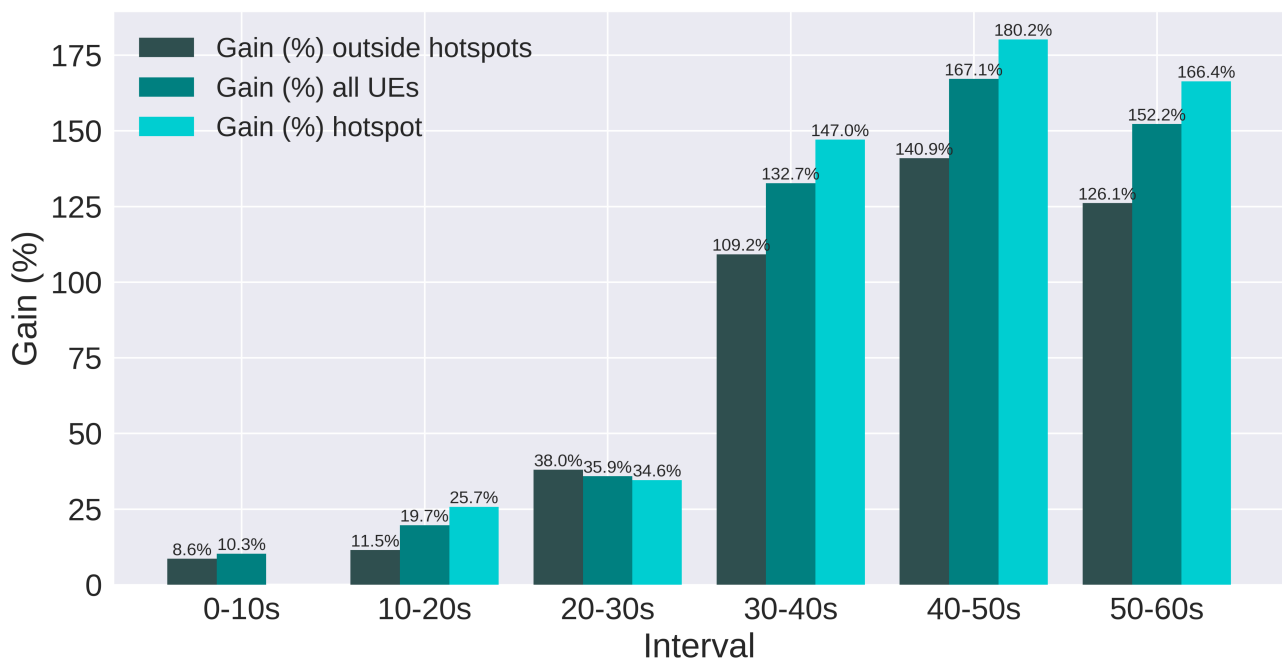


Figure 18. Scenario 2 of the proposed solution: SINR Gain (%) for each group of users.

Furthermore, Scenario 2 investigates the algorithm's performance with more users allocated in the hotspots and with different configurations for the *BandwidthDistribution* parameter, which also has a significant impact on the evaluated scenario (Section 8). This scenario allocates less bandwidth for the common sub-band and more bandwidth for the private sub-band.

Figure 18 presents the SINR gain for Scenario 2. Similar to Scenario 1, all groups of users have significant gain, and it is approximately doubled for the last three intervals, compared to the previous scenario. This is mainly due to the double amount of users on each hotspot of Scenario 2, which degrades the SINR more severely. Since the SINR is the target metric of the proposed algorithm, the gain is also doubled, showing that the QL is efficient in improving the SINR, even if the sub-bands are configured differently.

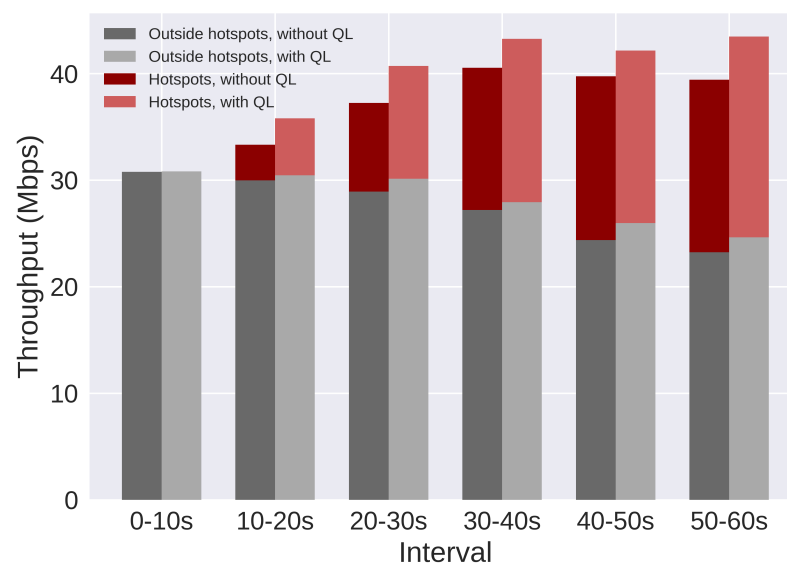


Figure 19. Scenario 1 of the proposed solution: throughput results for each group of users, with and without Q-Learning.

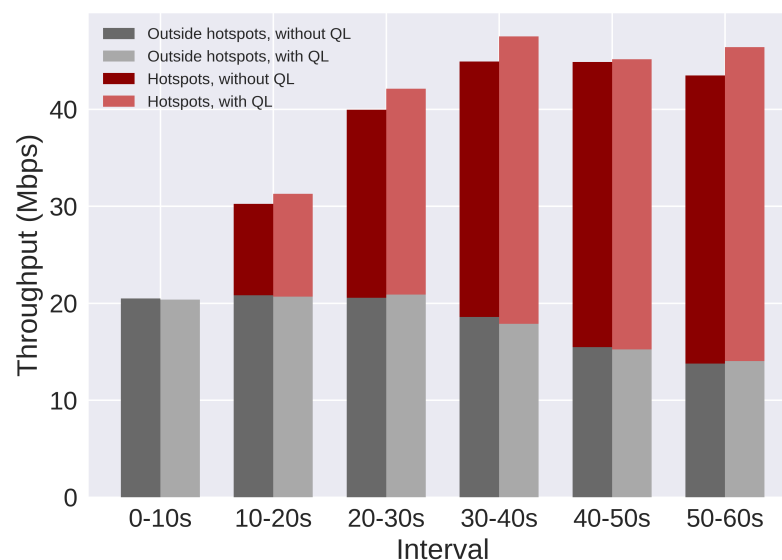


Figure 20. Scenario 2 of the proposed solution: throughput results for each group of users, with and without Q-Learning.

Figure 20 presents throughput results for Scenario 2. There is a relevant gain for the average throughput and for hotspot users. However, the gain is minimal to users outside

of hotspots. On the third interval (30 to 40 s), these users suffer a small loss, but the gain from hotspot users compensates, resulting in a gain in the system average throughput.

Summarizing, the throughput percentage gain is presented in Figures 21 and 22, for scenarios 1 and 2, respectively. The smallest gain obtained for hotspot users was 5.4% in scenario 1 and 1.7% in scenario 2. On the other hand, the highest gain for these users was 59.8% in scenario 1 and 12.5% in scenario 2. Users outside the hotspots presented loss in some intervals for scenario 2. Nevertheless, all intervals presented gain for hotspot users. These users were not present in the plots for the first interval because there were no active hotspots users. The worst and best percentage gain for each scenario and group of users are summarized in Table 12.

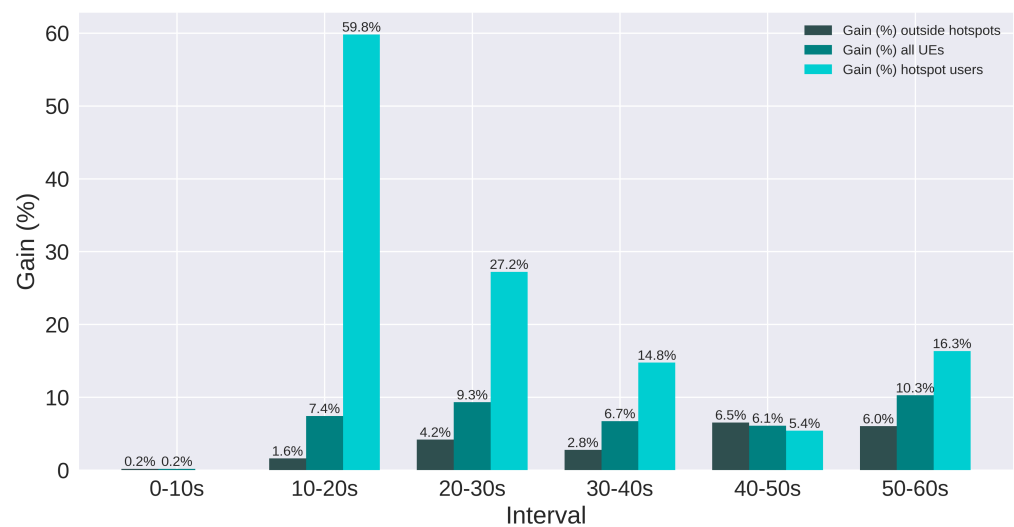


Figure 21. Scenario 1 of the proposed solution: throughput gain (%) for each group of users.

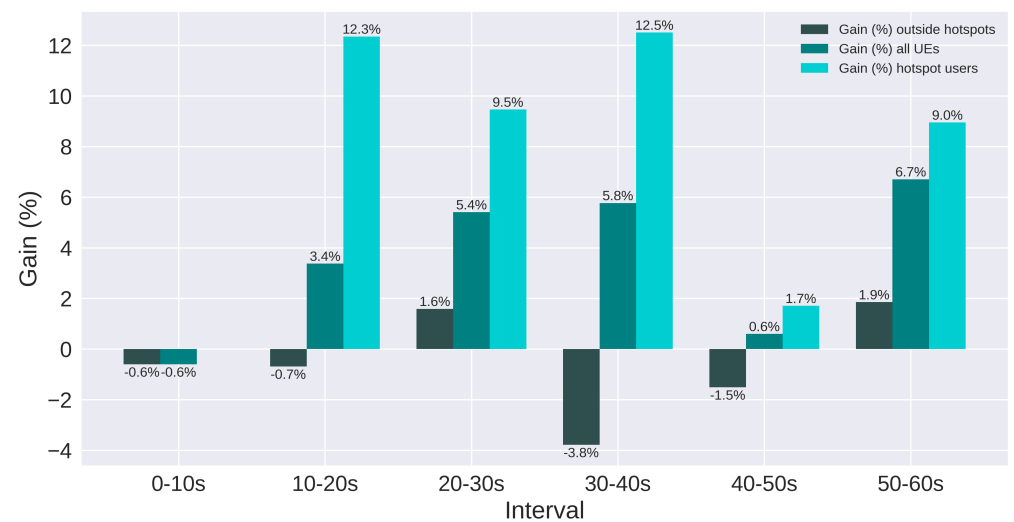


Figure 22. Scenario 2 of the proposed solution: throughput gain (%) for each group of users.

Table 12. Best and worst results for SINR and throughput of the Q-Learning solution, for each group of users.

	Best SINR Gain	Worst SINR Gain	Best Tput Gain	Worst Tput Gain
Scenario 1				
Hotspot users	94%	35.3%	59.8%	5.4%
Outside hotspots	99.1%	11.1%	6.5%	0.2%
All Users	96.7%	12.8%	10.3%	0.2%
Scenario 2				
Hotspot users	180.2%	25.7%	12.5%	1.7%
Outside hotspots	140.9%	8.6%	1.9%	−3.8%
All Users	167.1%	10.3%	6.7%	−0.6%

11. Conclusions

In recent years, mobile networks have experienced an increase in the number of scenarios with challenging requirements, especially due to rapid growth in traffic demand and the arrival of new services, such as eMBB. Likewise, user behavior has changed, creating scenarios that are hard to predict. This paper discussed the challenges related to a dynamic hotspot scenario and how to mitigate the resulting performance loss. A solution was presented to coordinate ICI using FFR-based techniques. The algorithm dynamically regulates the Strict FR parameters, using Q-Learning (QL).

The results indicate that the solution effectively optimizes the SINR in the evaluated scenarios, which also resulted in throughput gain, even though this relationship is not always true. Depending on the scenario, the SINR gain may only result in a decrease of the PLR, which is also important for the overall performance of the system.

In some cases, the gain was higher for hotspot users, even though the algorithm was not configured to act based on their performance. Besides, even in these cases, the algorithm also improved the overall system performance. The results also indicate a tendency of greater SINR gain when the scenario suffers with more interference. The higher SINR gain obtained was 180% for hotspot users, 167.1% for average SINR, and 140% for users outside hotspots.

12. Future Works

The development of research usually involves various choices in order to define the scope of the paper. As a consequence, some unexplored possibilities and possible improvements are listed below as an encouragement to future endeavors.

- Use other metrics to define the QL states and actions, such as the BandwidthDistribution. As presented in Section 10.3, these metrics can also be the combination of different variables. As a result, the algorithm is expected to improve in flexibility and efficiency when using two parameters, adapting to a broader set of scenarios. For example, the throughput could be included as part of the reward;
- Expand proposed scenarios: vary number of users, add mobility, varying number and location of hotspots (which could also appear in random locations), and test different values for BandwidthDistribution;
- Evaluate the system using different metrics, such as the relation between convergence speed, the amount of state/action pairs, or the Packet Loss Ratio (PLR);
- Apply the solution on a system without isotropic antennas, such that transmission is made in sectors within a cell;
- Apply the presented solution for 5G New Radio (NR) or the LTE UL, given that the interference in the UL has different characteristics, when compared to DL;
- Provide a similar solution, replacing the RL algorithm for, e.g., multi-armed bandit, providing a simpler solution.

Author Contributions: Conceptualization, I.D.d.R., V.A.d.S.J.; data curation, I.D.d.R.; formal analysis, I.D.d.R.; supervision, V.A.d.S.J.; visualization, I.D.d.R.; writing—original draft, I.D.d.R.; writing—reviewing and editing, V.A.d.S.J. All authors have read and agreed to the published version of the manuscript.

Funding: This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES)—Finance Code 001.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The proof of concept simulations provided by this paper was supported by the High Performance Computing Center at UFRN (NPAD/UFRN).

Conflicts of Interest: The authors declare that they have no competing interests.

References

1. Cisco; Cisco Systems, I. Cisco Visual Networking Index: Forecast and Trends, 2017–2022. *Cisco White Paper* **2019**.
2. GSMA Association. *The Mobile Economy, 2019*; GSMA Association: London, UK, 2019.
3. Shannon, C. Communication In The Presence Of Noise (Republished). *Proc. IEEE* **1998**, *86*, 447–457. [\[CrossRef\]](#)
4. Zheng, K.; Li, Y.; Zhang, Y.; Jiang, Z.; Long, H. Performance analysis and evaluation of deployment in small cell networks. *KSII Trans. Internet Inf. Syst.* **2015**, *9*, 886–900. [\[CrossRef\]](#)
5. Cavalcanti, F.R.P.; Maciel, T.F.; Freitas, W.C., Jr.; Silva, Y.C.B. *Comunicação Móvel Celular*, 1st ed.; Elsevier: Rio de Janeiro, Brazil, 2018.
6. Diggavi, S.N.; Al-dhahir, N.; Member, S. Great Expectations: The Value of Spatial. *Wirel. Netw.* **2004**, *92*, 219–270.
7. Adhikary, A.; Nam, J.; Ahn, J.Y.; Caire, G. Joint spatial division and multiplexing-The large-scale array regime. *IEEE Trans. Inf. Theory* **2013**, *59*, 6441–6463. [\[CrossRef\]](#)
8. Adhikary, A.; Dhillon, H.S.; Caire, G. Massive-MIMO meets HetNet: Interference coordination through spatial blanking. *IEEE J. Sel. Areas Commun.* **2015**, *33*, 1171–1186. [\[CrossRef\]](#)
9. Muirhead, D.; Imran, M.A.; Arshad, K. A Survey of the Challenges, Opportunities and Use of Multiple Antennas in Current and Future 5G Small Cell Base Stations. *IEEE Access* **2016**, *4*, 2952–2964. [\[CrossRef\]](#)
10. Yaacoub, E.; Husseini, M.; Ghaziri, H. An overview of research topics and challenges for 5G massive MIMO antennas. In Proceedings of the 2016 IEEE Middle East Conference on Antennas and Propagation (MECAP), Beirut, Lebanon, 20–22 September 2016; pp. 1–4. [\[CrossRef\]](#)
11. Khandekar, A.; Bhushan, N.; Tingfang, J.; Vanghi, V. LTE-Advanced: Heterogeneous Networks. In Proceedings of the 2010 European Wireless Conference (EW), Lucca, Italy, 12–15 April 2010.
12. Damjanovic, A.; Montojo, J.; Wei, Y.; Ji, T.; Luo, T.; Vajapeyam, M.; Yoo, T.; Song, O.; Malladi, D. A survey on 3GPP heterogeneous networks. *IEEE Wirel. Commun.* **2011**, *18*, 10–21. [\[CrossRef\]](#)
13. Erlinghagen, K.; Dusza, B.; Wietfeld, C. Dynamic Cell Size Adaptation and Inter-cell Interference Coordination in LTE HetNets. In Proceedings of the 2013 IEEE 78th Vehicular Technology Conference (VTC Fall), Las Vegas, NV, USA, 2–5 September 2013.
14. Chandrasekhar, V.; Andrews, J.G.; Gatherer, A. Femtocell Networks: A Survey. *IEEE Commun. Mag.* **2008**, *46*, 59–67. [\[CrossRef\]](#)
15. Al-Turjman, F.; Ever, E.; Zahmatkesh, H. Small cells in the forthcoming 5G/IoT: Traffic modelling and deployment overview. *IEEE Commun. Surv. Tutorials* **2019**, *21*, 28–65. [\[CrossRef\]](#)
16. Mudassir, A.; Akhtar, S.; Kamel, H. Survey on Inter-cell Interference Coordination in LTE-Advanced Heterogeneous Networks. In Proceedings of the 2016 Sixth International Conference on Innovative Computing Technology (INTECH), Dublin, Ireland, 24–26 August 2016.
17. Ree, M.D.; Mantas, G.; Radwan, A.; Mumtaz, S.; Rodriguez, J.; Otung, I.E. Key Management for Beyond 5G Mobile Small Cells: A Survey. *IEEE Access* **2019**, *7*, 59200–59236. [\[CrossRef\]](#)
18. Chang, B.J.; Liou, S.H.; Liang, Y.H. Cooperative communication in ultra-dense small cells toward 5G cellular communication. In Proceedings of the 2017 8th IEEE Annual Information Technology, Electronics and Mobile Communication Conference, IEMCON, Vancouver, BC, Canada, 3–5 October 2017; pp. 365–371. [\[CrossRef\]](#)
19. Gawlowicz, P.; Baldo, N.; Miozzo, M. An Extension of the ns-3 LTE Module to Simulate Fractional Frequency Reuse Algorithms. In Proceedings of the Workshop on ns-3-WNS3 2015, Barcelona, Spain, 13–14 May 2015.
20. Hamza, A.; Khalifa, S.; Hamza, H.; Elsayed, K. A Survey on Inter-Cell Interference Coordination Techniques in OFDMA-Based Cellular Networks. *IEEE Commun. Surv. Tutorials* **2013**, *15*, 1642–1670. [\[CrossRef\]](#)
21. Kimura, D.; Seki, H. Inter-Cell Interference Coordination (ICIC) Technology. *Fujitsu Sci. Tech. J.* **2012**, *48*, 89–94.
22. Xie, Z.; Walke, B. Enhanced Fractional Frequency Reuse to Increase Capacity of OFDMA Systems. In Proceedings of the 3rd International on New Technologies, Mobility and Security (NTMS 2009), Cairo, Egypt, 20–23 December 2009.
23. Sesia, S.; Toufik, I.; Baker, M. *LTE—The UMTS Long Term Evolution: From Theory to Practice*, 2nd ed.; Wiley: Chichester, UK 2011.

24. Holma, H.; Toskala, A. *LTE Advanced: 3GPP Solution for IMT-Advanced*, 1st ed.; Wiley: Chichester, UK, 2012.
25. 3GPP. *TS 36.213: Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Layer Procedures (Release 8)*; Technical Report; 3GPP: Valbonne, France, 2009.
26. 3GPP. *5G; NR; Overall Description; Technical Specification 38.300 version 15.3.1 Release 15*; Technical Report; 3GPP: Valbonne, France, 2018.
27. Soret, B.; De Domenico, A.; Bazzi, S.; Mahmood, N.H.; Pedersen, K.I. Interference Coordination for 5G New Radio. *IEEE Wirel. Commun.* **2018**, *25*, 131–137. [[CrossRef](#)]
28. Singh, R.; Siva Ram Murthy, C. Techniques for Interference Mitigation Using Cooperative Resource Partitioning in Multitier LTE HetNets. *IEEE Syst. J.* **2018**, *12*, 843–853. [[CrossRef](#)]
29. López-Pérez, D.; Ding, M.; Claussen, H.; Jafari, A.H. Towards 1 Gbps/UE in Cellular Systems: Understanding Ultra-Dense Small Cell Deployments. *IEEE Commun. Surv. Tutorials* **2015**, *17*, 2078–2101. [[CrossRef](#)]
30. Shirakabe, M.; Morimoto, A.; Miki, N. Performance evaluation of inter-cell interference coordination and cell range expansion in heterogeneous networks for LTE-Advanced downlink. In Proceedings of the 2011 8th International Symposium on Wireless Communication Systems, Aachen, Germany, 6–9 November 2011; pp. 844–848. [[CrossRef](#)]
31. Shen, L.H.; Feng, K.T. Joint beam and subband resource allocation with QoS requirement for millimeter wave MIMO systems. In Proceedings of the 2017 IEEE Wireless Communications and Networking Conference (WCNC), San Francisco, CA, USA, 19–22 March 2017; pp. 1–6. [[CrossRef](#)]
32. Rüegg, T.; Hassan, Y.; Wittneben, A. User cooperation enabled traffic offloading in urban hotspots. In Proceedings of the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC, Valencia, Spain, 4–8 September 2016. [[CrossRef](#)]
33. Khan, S.A.; Colak, S.A.; Kavak, A.; Kucuk, K. A User Location Distribution Based FFR Strategy for Efficient Utilization of Radio Resources in LTE-A HetNets. In Proceedings of the 1st International Informatics and Software Engineering Conference: Innovative Technologies for Digital Transformation, IISEC 2019-Proceedings, Ankara, Turkey, 6–7 November 2019. [[CrossRef](#)]
34. Li, X.; Liu, Z.; Qin, N.; Jin, S. FFR Based Joint 3D Beamforming Interference Coordination for Multi-Cell FD-MIMO Downlink Transmission Systems. *IEEE Trans. Veh. Technol.* **2020**, *69*, 3105–3118. [[CrossRef](#)]
35. Zheng, C.; Hailin, Z. A Quasi-Perfect Resource Allocation Scheme for Optimizing the Performance of Cell-Edge Users in FFR-Aided LTE-A Multicell Networks. *IEEE Commun. Lett.* **2019**, *23*, 918–921. [[CrossRef](#)]
36. Mitsolidou, C.; Vagionas, C.; Mesodiakaki, A.; Maniotis, P.; Kalfas, G.; Roeloffzen, C.G.; van Dijk, P.W.; Oldenbeuving, R.M.; Miliou, A.; Pleros, N. A 5G C-RAN optical fronthaul architecture for hotspot areas using OFDM-based analog IFoF waveforms. *Appl. Sci.* **2019**, *9*, 4059. [[CrossRef](#)]
37. Ma, B.; Yang, B.; Zhu, Y.; Zhang, J. Context-aware proactive 5g load balancing and optimization for urban areas. *IEEE Access* **2020**, *8*, 8405–8417. [[CrossRef](#)]
38. Abdullahi, S.U.; Liu, J.; Mohadeskasaei, S.A. Analytical evaluation of FFR-aided heterogeneous cellular networks with optimal double threshold. *KSII Trans. Internet Inf. Syst.* **2017**, *11*, 3370–3392. [[CrossRef](#)]
39. ns-3. ns-3 Documentation. Página na Internet, ns-3. 2020. Available online: <https://www.nsnam.org/> (accessed on 24 October 2021).
40. ns-3. Design Documentation. Página na Internet, ns-3. 2019. Available online: <https://www.nsnam.org/> (accessed on 24 October 2021).
41. Hamza, A.S.; Khalifa, S.S.; Hamza, H.S.; Elsayed, K. A Survey on Inter-Cell Interference Co-ordination Techniques of LTE Networks. *IEEE Commun. Surv. Tutor.* **2013**, *4*, 1–29.
42. 3GPP. *TS 36.133: Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Requirements for Support of Radio Resource Management (Release 8)*; Technical Report; 3GPP: Valbonne, France, 2010.
43. de Santana, P.M.; de Sousa, V.A., Jr.; Abinader, F.M., Jr.; de C. Neto, J.M. DM-CSAT: A LTE-U/Wi-Fi coexistence solution based on reinforcement learning. *Telecommun. Syst.* **2019**, *1*, 1–12. [[CrossRef](#)]
44. Montgomery, D.C. *Estatística Aplicada e Probabilidade Para Engenheiros*, 4th ed.; ETC: Leusden, The Netherlands, 2009.
45. Mathematics, E. *Chi-Squared Distribution*; Web page; Springer: Berlin/Heidelberg, Germany, 2020.
46. Mathematics, E. *Fisher-F-Distribution*; Web page; Springer: Berlin/Heidelberg, Germany, 2020.
47. Trejo Narváez, O.A.; Miramá Pérez, V.F. Machine learning algorithms for inter-cell interference coordination. *Sist. Telemática* **2018**, *16*, 37–57. [[CrossRef](#)]
48. Haykin, S.S. *Neural Networks and Learning Machines*, 3rd ed.; Pearson Education: Upper Saddle River, NJ, USA, 2009.
49. Richard, S.; Sutton, A.G.B. *Reinforcement Learning: An Introduction*; MIT Press: Cambridge, MA, USA, 2018.
50. Chen, Z.; Chuai, G.; Gao, W. A novel ICIC scheme for reducing intra cluster interference in LTE-Hi network. In Proceedings of the 2016 IEEE International Conference on Network Infrastructure and Digital Content (IC-NIDC), Beijing, China, 23–25 September 2017; Volume 129, pp. 262–267. [[CrossRef](#)]
51. Simsek, M.; Bennis, M.; Czylwik, A. Dynamic Inter-Cell Interference Coordination in HetNets: A reinforcement learning approach. In Proceedings of the 2012 IEEE Global Communications Conference (GLOBECOM), Anaheim, CA, USA, 3–7 December 2012; pp. 5446–5450. [[CrossRef](#)]
52. Morozs, N.; Clarke, T.; Grace, D. Distributed Heuristically Accelerated Q-Learning for Robust Cognitive Spectrum Management in LTE Cellular Systems. *IEEE Trans. Mob. Comput.* **2016**, *15*, 817–825. [[CrossRef](#)]

-
53. de Santana, P.M.; de Sousa, V.A., Jr.; Abinader, F.M., Jr.; de C. Neto, J.M. GTDM-CSAT: An LTE-U self Coexistence Solution based on Game Theory and Reinforcement Learning. *J. Commun. Inf. Syst. (JCIS)* **2019**, *34*, 169–177. [[CrossRef](#)]
 54. Watkins, C.; Dayan, P. Q-learning. *Mach. Learn.* **1992**, *8*, 279–292. [[CrossRef](#)]
 55. Watkins, C. Learning From Delayed Rewards. Ph.D. Thesis, King's College, London, UK, 1989.
 56. Bazzo, J.J.; de Sousa, V.A., Jr. Method and Apparatus for Admission Control and Forced Handover in a Multi-Layer Network Configuration. U.S. Patent US8660086B2, 11 October 2011.