Departament de Teoria
del Senyal i Comunicacions
UNIVERSITAT POLITÈCNICA DE CATALUNYA

# Contribution to Dynamic Spectrum Assignment in Multicell OFDMA Networks

## *A self-organization and machine learning approach*

**DISSERTATION**

**Doctor of Philosophy**

**Francisco Bernardo Álvarez**

This dissertation is submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Doctorat de Teoria del Senyal i Comunicacions in the Departament de Teoria del Senyal i Comunicacions (TSC) of the Universitat Politècnica de Catalunya (UPC).

**Thesis Advisor**

Ramon Agustí Comes

Barcelona, April 2010

*A mi familia, a mis amigos y a Cristina.*

# Abstract

Next fourth generation (4G) of cellular mobile networks envisage a radio interface based on OFDMA (Orthogonal Frequency Division Multiple Access). OFDMA offers frequency diversity and robustness against multipath channel propagation thanks to the division of a wide bandwidth into small OFDMA frequency resources, so that an efficient spectrum usage is attained. However, one important challenge in a 4G OFDMA-based radio interface of a cellular network is the way in which OFDMA frequency resources are assigned to cells. First, intercell interference must be mitigated to achieve the highest spectral efficiency. Second, traffic loads could vary along time and space, so typical fixed spectrum assignment patterns (i.e., frequency planning) could lead to lack of spectrum resources in some cells or underutilization of them in others. Third, future regulatory spectrum frameworks will change the mindset about the usage of the spectrum by planning the co-existence of primary (licensees) and secondary users of the spectrum in the same geographical area. Hence, an adequate primary management of the spectrum could ease the appearance of spectrum usage opportunities for secondary users at the same time that the primary operator could obtain a new revenue income for that usage. Finally, future cellular scenarios will tend to be decentralized, especially with the appearance of new femtocell deployments (short-range user-deployed access points in the operator's licensed spectrum band), where a high degree of independency when deciding the usage of OFDMA frequency resources will be needed, since, obviously, centralized frequency planning tasks has little practical sense in those scenarios.

This thesis contributes to the research on the spectrum assignment in 4G OFDMA-based mobile cellular networks by proposing a solution to dynamically manage the cell-by-cell spectrum assignment. To this end, adequate Dynamic Spectrum Assignment (DSA) strategies and a practical framework to execute them are proposed. In order to reduce operational costs and to reduce human intervention, the DSA framework has been designed based on self-organization so that the network is able to autonomously (*i*) observe the performance of current spectrum assignment, (*ii*) analyze if a new spectrum assignment is needed, and (*iii*) decide a new spectrum assignment to better adapt to networks conditions. Furthermore, a centralized and decentralized functional architecture is proposed so that the framework can be applied to a vast number of scenarios, ranging from typical macrocell scenarios where centralized control is employed, to future femtocell scenarios where nodes are almost independent and require autonomous spectrum assignment decisions at the cell level. The decision task of the framework resides on proposed DSA strategies, where one of them is based on machine learning to exploit knowledge previously acquired in the past. Moreover, this machine learning strategy tends to select a spectrum assignment that is optimal in the sense that maximizes a given reward signal appropriately defined in terms of network performance metrics (e.g., spectral efficiency, SINR, among others). Certainly, self-organization and machine learning on the context of spectrum assignment in OFDMA based radio interfaces have been little exploited and thus constitutes a major novelty of the work of this thesis.

Performance results reveal important improvements over state-of-the-art strategies in terms of spectral efficiency (i.e. in bits/s/Hz), users' QoS satisfaction, fairness between the throughput obtained by users, and capacity for generating opportunities for secondary spectrum usage in large geographical areas. Also, the proposed DSA framework, based on self-organization, demonstrates appealing capabilities from the perspective of initial deployment, where nodes are able to self-configure after switch-on introducing a minimal impact on the already deployed system, being then a practical contribution to solve the deployment of thousands of femtocells in macrocell environments.

# Resumen

La próxima cuarta generación (4G) de redes de comunicaciones móviles celulares considera una interfaz radio basada en OFDMA (Orthogonal Frequency Division Multiple Access). Esta tecnología ofrece robustez a la propagación multicamino y diversidad en frecuencia gracias a la división del ancho de banda de operación en un conjunto de pequeños subcanales en frecuencia para así alcanzar un uso eficiente del espectro. Sin embargo, un importante desafío en una interfaz radio OFDMA es la manera en la que los subcanales en frecuencia se asignan a las distintas celdas. En primer lugar, la carga de tráfico podría variar a lo largo del tiempo y del espacio, de modo que los clásicos patrones fijos de asignación de espectro (es decir, los esquemas de planificación de frecuencias) pueden conducir a una carencia de recursos en ciertas celdas o a una falta de aprovechamiento de los mismos en otras. En segundo lugar, los futuros marcos reguladores del espectro cambiarán la mentalidad sobre su uso, planteando la coexistencia de usuarios primarios y secundarios del espectro en una misma área geográfica. Por lo tanto, una adecuada gestión del espectro primario podría facilitar la aparición de oportunidades del uso del espectro para usuarios secundarios a la vez que el operador podría obtener una nueva entrada de ingresos por ese uso. Finalmente, los futuros escenarios celulares tenderán a ser descentralizados, especialmente con la aparición de nuevos despliegues basados en *femtoceldas* (puntos de acceso de limitada cobertura desplegados por el propio usuario en la banda espectral en la que el operador tiene licencia), donde se requerirá un alto grado de independencia a la hora de decidir los canales que usa cada femtocelda, ya que, obviamente, las tareas centralizadas planificación de frecuencias tienen poco sentido práctico en estos escenarios.

Esta tesis contribuye a la investigación sobre la asignación de espectro en redes móviles celulares 4G basadas en OFDMA proponiendo una solución para manejar dinámicamente la asignación de espectro por celda. Con este fin, se proponen estrategias dinámicas asignación de espectro (en inglés *Dynamic Spectrum Assigment*: DSA) y un marco práctico para ejecutarlas. Para reducir costes operacionales y la intervención humana en el proceso, el marco DSA propuesto se ha diseñado basándose en conceptos de autoorganización de modo que la red puede de forma autónoma (*i*) observar el funcionamiento de la asignación actual de espectro, (*ii*) analizar si una nueva asignación de espectro es necesaria, y (*iii*) decidir una nueva asignación de espectro que se adapte mejor a las condiciones de la red. Además, se propone una arquitectura funcional centralizada y descentralizada que permite que el marco DSA pueda aplicarse a varios escenarios, desde escenarios macrocelulares donde típicamente se emplea un control centralizado, a futuros escenarios con femtoceldas donde los nodos son prácticamente independientes y requieren de decisiones autónomas a nivel de celda para la asignación de espectro. La tarea de decisión del marco DSA reside en las estrategias DSA propuestas, donde una de ellas se basa en el aprendizaje máquina para explotar el conocimiento adquirido previamente en el pasado. Además, esta estrategia tiende a seleccionar una asignación de espectro óptima en el sentido de que maximiza una señal de recompensa definida apropiadamente en términos de métricas del funcionamiento de la red (e.g., eficiencia espectral, SINR, entre otras). Ciertamente, la autoorganización y el aprendizaje máquina en el contexto de la asignación de espectro en interfaces radio basadas en OFDMA se han explotado poco y constituyen así una novedad importante derivada del trabajo de esta tesis.

Los resultados revelan importantes mejoras sobre estrategias del estado del arte en términos de eficiencia espectral (en bits/s/Hz), satisfacción de la calidad de servicio de los usuarios, *fairness* entre el throughput obtenido por los usuarios, y la capacidad para generar oportunidades para el uso secundario

del espectro en grandes áreas geográficas. También, el marco DSA propuesto basado en autoorganización muestra atractivas capacidades desde la perspectiva del despliegue inicial, donde los nodos son capaces de autoconfigurarse después de su encendido introduciendo un impacto mínimo en sistema ya desplegado. Así el marco propuesto constituye una contribución práctica para solucionar el despliegue de millares de femtoceldas en un escenario macrocelular.

# Acknowledgements

First of all, I would like to deeply thank my thesis advisor Prof. Ramon Agustí for his guidance, ideas and constant aid during the work of this thesis. Definitely, his advice has made me maturing professionally and personally. Also, I would like to show my most sincere gratitude to Dr. Oriol Sallent and Dr. Jordi Pérez-Romero for their valuable support and fruitful comments during this research, which, for sure have help me to improve this work.

Moreover, the work of this thesis has been developed under the FPU (Formación de Profesorado Universitario) grants of the by the Spanish Education Council whose financial support is hereby acknowledged.

In addition, I would like to appreciate the assistance of Prof. Ferran Casadevall and Dr. Anna Umbert during my first research activities at the Grup de Recerca de Comunicacions Mòbils (GRCM). Also, I am very thankful to all the colleagues I have had the opportunity to meet at the GRCM since working close to them has certainly eased the way to complete this thesis. Especially thanks to Nemanja, Jad, Xavi, Miguel, Hiram and Vuk. Besides, I am also grateful to my flat mate Joan Tomàs who has made my stay at Barcelona these four years as if I was at home. Also, I would like to extend my gratitude to colleagues of the Departamento de Teoría de la Señal y Comunicaciones (DTSC) of the Universidad de Sevilla for the good moments we have shared during the last part of this work.

Moving towards more personal acknowledgements, I would like to thank my friends Antonio, Jose and Miky who, even in the distance, have showed a close friendship and the conviction that I could, some day, manage to finish this work. Also thanks to Paqui and Manolo for their warm support. Noneless to say that I am particularly indebted to my family for being always to my side in my decisions and their monumental and unconditional support. Last, but not least, I want to profoundly thank Cristina for her enormous patient and continuous encouragement, and, for the precious gift that she has given me by sharing these years with me.

# Contents

# List of Figures

# List of Tables

# List of Abbreviations

| | |
|---|---|
| **2G** | Second generation |
| **3G** | Third generation |
| **3GPP** | 3rd Generation Partnership Project |
| **4G** | Fourth Generation |
| **A/D** | Analogical/Digital |
| **AMC** | Adaptive Modulation and Coding |
| **ARQ** | Automatic Repeat Request |
| **B3G** | Beyond 3G |
| **BER** | Bit Error Rate |
| **BLU** | Bernoulli Logistic Unit |
| **CAPEX** | Capital Expenditure |
| **CCE** | Cell Characterization Entity |
| **CDF** | Cumulative Distribution Function |
| **CDMA** | Code Division Multiple Access |
| **CP** | Cyclic Prefix |
| **CQI** | Channel Quality Indicator |
| **CSI** | Channel State Information |
| **D/A** | Digital/Analogical |
| **DCA** | Dynamic Channel Assignment |
| **DFT** | Discrete Fourier Transform |
| **DL** | Downlink (forward link) |
| **DP** | Dynamic Programming |
| **DSA** | Dynamic Spectrum Assignment |
| **DCA** | Dynamic Channel Assignment |
| **DSL** | Digital Subscriber Link |
| **DVB-T** | Digital Video Broadcasting - Terrestrial |
| **eNB** | Evolved Node B |
| **EPC** | Evolved Packet Core |
| **E-UTRAN** | Evolved UTRAN |
| **FCA** | Fixed Channel Assignment |
| **FDD** | Frequency Division Duplex |
| **FDMA** | Frequency Division Multiple Access |
| **FFT** | Fast Fourier Transform |
| **FRF** | Frequency Reuse Factor |
| **GERAN** | GSM Edge Radio Access Network |
| **GPRS** | General Packet Radio System |

| | |
|---|---|
| **GSM** | Global System for Mobile communications |
| **HARQ** | Hybrid Automatic Repeat Request |
| **HEUR-DSA** | Heuristic DSA |
| **HSPA** | High Speed Packet Access |
| **ICI** | Inter-Carrier Interference |
| **ICIC** | Intercell Interference Coordination |
| **IDFT** | Inverse Discrete Fourier Transform |
| **IEEE** | Institute of Electrical and Electronic Engineers |
| **IFFT** | Inverse Fast Fourier Transform |
| **IP** | Internet Protocol |
| **ISI** | Inter-Symbol Interference |
| **JRRM** | Joint Radio Resource Management |
| **KPI** | Key Performance Indicator |
| **LTE** | Long Term Evolution |
| **MDP** | Markov Decision Process |
| **MIMO** | Multiple Input Multiple Output |
| **MME** | Mobility Management Entity |
| **MWM** | Multi Wall Model |
| **NCE** | Network Characterization Entity |
| **NTB** | Normalized Throughput Bound |
| **OAM** | Operation and Maintenance |
| **OFDM** | Orthogonal Frequency Division Multiplexing |
| **OFDMA** | Orthogonal Frequency Division Multiple Access |
| **OPEX** | Operational Expenditure |
| **PAPR** | Peak-to-Average Power Ratio |
| **PDF** | Probability Density Function |
| **PF** | Proportional Fair |
| **P-GW** | PDN Gateway |
| **PHY** | Physical Layer |
| **PL** | Path Loss |
| **PR** | Partial frequency Reuse |
| **PUSC** | Partial Usage of Subcarriers |
| **QAM** | Quadrature Amplitude Modulation |
| **QCI** | QoS Class Identifier |
| **QoS** | Quality of Service |
| **QPSK** | Quadrature Phase Shift Keying |
| **RAT** | Radio Access Technology |

| | |
|---|---|
| **RB** | Resource Block |
| **RBC** | Radio Bearer Control |
| **REINFORCE** | REward Increment = Nonnegative Factor times Offset Reinforcement times Characteristic Eligibility |
| **RL** | Reinforcement Learning |
| **RL-DSA** | Reinforcement Learning DSA |
| **RMSE** | Root Mean Square Error |
| **RNTP** | Relative Narrowband Transmit Power |
| **RR** | Round Robin |
| **RRM** | Radio Resource Management |
| **RX** | Reception |
| **SC-FDMA** | Single Carrier Frequency Division Multiple Access |
| **SGW** | Server Gateway |
| **SINR** | Signal to Interference plus Noise Ratio |
| **SIR** | Signal to Interference Ratio |
| **SNR** | Signal to Noise Ratio |
| **SON** | Self-Organizing Networks |
| **SR** | Soft frequency Reuse |
| **SS** | Subscriber Station |
| **STS** | Short-Term Scheduler |
| **TD** | Temporal Difference learning |
| **TDD** | Time Duplex Division |
| **TDMA** | Time Division Multiple Access |
| **TV** | Television |
| **TX** | Transmission |
| **UE** | User Equipment |
| **UL** | Uplink (reverse link) |
| **UMTS** | Universal Mobile Telecommunications System |
| **URS** | Useful Released Surface |
| **UTRAN** | Universal Terrestrial Radio Access Network |
| **VoIP** | Voice over IP |
| **WiMAX** | Worldwide Interoperability for Microwave Access |

# Glossary

**Action-Selection Probability** The probability that a given reinforcement learning agent selects a given action.

**Cellular Mobile Network** A wireless infrastructure where service is provided through the deployment of several access points or base stations that cover a given geographical area called cell.

**Channel-Aware Scheduling** A packet scheduling that takes into account channel status to take its decisions.

**Chunk** Radio resource in frequency domain in an OFDMA resource grid, composed by a group of contiguous OFDM subcarriers.

**Cyclic Prefix** OFDM technique where the final part of OFDM symbols is copied at the beginning to combat intercarrier and intersymbol interference.

**Dynamic Spectrum Assignment** A cell-by-cell spectrum assignment that is changed during the cellular network operation to adapt to variable network conditions.

**Femtocell** A low-range, user-deployed access point to mainly provide indoor coverage at homes or offices.

**Frame** A radio resource in the time domain in an OFDMA resource grid.

**Frequency Planning** The task of assigning frequency resources to cells prior to the operation of the cellular network. This assignment usually remains unaltered for very long periods of time.

**Frequency Reuse Factor** A spectrum assignment in a cellular deployment where the available frequency resources are equally divided into several subgroups, where each one is assigned to a cell following a regular pattern.

**Intercell Interference** The interference between transmissions to/from users belonging to the different cells.

**Intracell Interference** The interference between transmissions to/from users belonging to the same cell.

**Key Performance Indicator** A metric that reveals significant information concerning state and functioning of a system and the quality of the delivered service.

**Learning Rule** The procedure employed by a reinforcement learning agent to modify action-selection probabilities after receiving reward.

| | |
|---|---|
| **Link Adaptation** | The task of dynamically assigning modulation and channel coding schemes to scheduled users' transmissions by packet scheduler. |
| **Macrocell** | A high-range, operator-deployed access point to mainly provide outdoor coverage. |
| **Measurement Report** | A message containing information about the metrics collected by a user from its serving cell or other cells. |
| **Medium-Term** | In this thesis, a period of time in the order of seconds, or tens of seconds. |
| **Multiuser Diversity** | In a multipath propagation environment, the property where different users experiment different realizations of the channel at the same time. |
| **OFDMA** | Multiple access technique where the radio interface is based on OFDM. |
| **Packet Scheduling** | The task of dynamically assigning OFDMA resource blocks to users' data transmissions. |
| **Partial frequency Reuse** | A spectrum assignment scheme where the inner and outer part of a cell employ different frequency reuse factors. Usually the outer part employs higher frequency reuse factors to mitigate intercell interference. |
| **Private Commons** | A spectrum regulatory framework where the primary (licensed) operators agree to open their spectrum to be used by a secondary spectrum market. |
| **Random Frequency Reuse Factor** | A spectrum assignment where cells randomly select a subset of the available frequency resources to operate. |
| **Reinforcement Learning** | A machine learning methodology where certain agents learn the best actions to take to maximize a given reward or payoff through a continuous interaction with an environment. |
| **Resource Block** | A specific chunk in a given frame in an OFDMA resource grid. |
| **Reward** | A scalar that measures the suitability of a given action taken by a reinforcement learning agent. |
| **Secondary Spectrum Market** | A way of exploiting a primary (licensed) spectrum by opportunistic, secondary users, which usually do not have a license to operate. |

| | |
|---|---|
| **Self-organization** | The ability of a set of agents or nodes to autonomously decide, though interaction with an environment local to each agent, certain actions to improve the global status of the whole set of agents. |
| **Self-organizing Network** | A cellular network that implements self-organization capabilities. |
| **Short-Term** | In this thesis, a period of time in the order of milliseconds. |
| **Soft frequency Reuse** | A spectrum assignment where the inner and outer part of a cell employ different powers per frequency resource. Inner resources are transmitted with lower power than outer resources. |
| **Spectral Efficiency** | Performance metric that measures the amount of successfully delivered bits per unit of time and spectrum. |
| **Spectrum Allocation** | The assignment of a block of spectrum by the spectrum regulator to be exploited for a given service provided by an operator. |
| **Spectrum Assignment** | A way of distributing frequency resources among cells. |
| **Throughput Fairness** | Performance metric that measures the balance between the throughput obtained by different users in a cell or group of cells. |
| **Two layer cellular deployment** | A cellular network where macrocells and femtocells coexist. |
| **Useful Released Surface** | Performance metric that measures the amount of spectrum that can be used by secondary users without harming primary transmissions in a given regional area. |
| **User's Throughput Dissatisfaction probability** | Performance metric that reveals the probability that the throughput obtained by a user in average is below a given target. |

# 1

# INTRODUCTION

## Outline

# 1 Introduction

Information society leads to a growing demand on accessing to information anywhere and anytime in a mobile world. Users expect to access to a wide range of applications, capabilities and services ubiquitously mostly using a single subscription. Thus, new value-added and dynamic services such as multimedia messages, podcasts, online gaming, Internet on the palm of a hand, or mobile TV have appeared, what at the same time has stimulated the development of more and more intelligent mobile terminals and network infrastructure. Certainly, a symbiosis between users, operators and manufacturers has been created, where, on the one hand, users require more frequently advanced services of higher quality and, on the other hand, operators and manufacturers aim at exploiting this increasing demand offering higher capacity networks and advanced devices. This results in a fast and constant development of the information society technologies where mobile communications have played and will continue playing a key role.

Regarding the advances in mobile communications in the last two decades, cellular networks have evolved from mostly circuit-switched 2G networks such as GSM (Global System for Mobile communications) offering service to an exponential growth of voice users in the 90s, to current packet-switched 3.5G networks such as HSPA (High Speed Packet Access) to cope with an increasing demand of data traffic motivated by the appearance of flat rates billing formulas for mobile Internet access. In fact, multimedia traffic is increasing far more rapidly than voice, and will increasingly dominate traffic flows, so that packet-switched services are gradually taking a predominant position over classical circuit-switched services.

As a result, future 4G networks envisaged for the next decade will continue with this tendency, adopting an all-IP (Internet Protocol) architecture, where classical circuit-switched voice services will be replaced by Voice-over-IP (VoIP) services. Besides, 4G networks go a step further in the development of current mobile networks by offering higher spectral efficiencies, lower latencies, lower power consumption, reduced costs per transmitted bit, and increased service provisioning (i.e., more services at lower cost with better user experience). As an initial step, 3GPP (3$^{rd}$ Generation Partnership Project) has defined the Long-Term Evolution (LTE) system, which constitutes an evolution of 3GPP standards in terms of radio interface and system architecture. On the other hand, from IEEE the WiMAX 802.16e standard has been defined as an initial pre-4G solution competing with LTE. Both approaches pave the way for LTE-Advanced and 802.16m respectively as full 4G systems that cope with IMT-Advanced requirements [1][2].

Therefore, one of the key requirements for future 4G networks is attaining the highest possible spectral efficiency on the radio interface. For instance, peak data rates of 100 Mbit/s and 1Gbit/s are expected in high and low mobility scenarios respectively [1]. To this end, 4G networks envisage the usage of the latest technical advances. Among others, Orthogonal Frequency Division Multiple Access (OFDMA), packet scheduling, Adaptive Modulation and Coding (AMC), and Multiple Input Multiple Output (MIMO) techniques emerge as the predominant technical solutions. OFDMA multicarrier transmission effectively combats typical mobile channel multipath propagation, provides frequency diversity in a broadband channel and allows for an efficient spectrum usage thanks to the division of a wide bandwidth into small frequency subchannels. Packet scheduling enables an efficient short-term temporal exploitation of OFDMA radio resources by dynamically assigning them to users taking into account the short-term channel quality for each user (i.e., selecting the user that best uses the channel).

AMC adjusts the modulation and coding rate for each scheduled transmission trying to achieve the highest spectral efficiency in bits/s/Hz. Finally, MIMO allows for transmission and reception over multiple antennas offering multiple parallel communication channels to (*i*) increase the number of users per cell if parallel channels are assigned to different users, (*ii*) increase a single user's data rate if parallel channels are assigned to the same user, or (*iii*) improve coverage if those parallel channels are used to provide spatial diversity through redundant communications.

However, some challenges in a mobile cellular deployment could arise as limiting factors for the achievement of high spectral efficiencies. For instance, high capacities per cell will be needed, what initially translates in an exhaustive reuse of the frequency band in the cellular deployment. This could lead to high intercell interference (i.e., the interference that two or more neighboring cells using the same frequency resource cause each other) considerably reducing the achievable spectral efficiency per OFDMA frequency resource.

On the other hand, the deployment of a 4G mobile network is likely to be in a scenario different to the one given in previous generations of mobile communications. Thus, new paradigms will impact the development next generation mobile networks as described in the following:

1.- *Femtocells*. It is expected that 4G cellular networks, as previous generations, will be composed of several layers in dense urban scenarios, where a layer of large coverage *macrocells* on the top of buildings is combined with other layers compound of small coverage *microcells* in streets or *picocells* inside buildings. However, a new type of cell site known as *femtocell* arises as a key paradigm that will take a very relevant role in the next decade. Femtocells are small range user-deployed base stations introduced at a considerable amount of random locations such as users' homes or offices [3], where end-to-end communication is given through a DSL (Digital Subscriber Line) link. Hence, from users' perspective they offer high data rates and improved coverage and service quality perception. On the other hand, from an operator's point of view, they increase capacity in terms of the number of users supported by the network. Also, femtocells extend indoor coverage avoiding that macrocell resources are destined to penetrate inside buildings where, typically, most of the traffic is generated. Finally, femtocells enable the exploitation of high frequency bands thanks to their short coverage requirements.

   However, femtocells introduce a big challenge in the way that the network is managed, because network operator has very limited control over the femtocells' positions and configuration (femtocells are property of the final user), and because a high number of deployed femtocells is expected (some studies in 2006 [4] foresaw about 32 millions of femtocells worldwide in 2011).

2.- *Self-organization*. 4G networks are being designed in a way that capital and especially operational expenditures reductions are achieved. In this context, Self-Organizing Networks (SON) concept [5] arises as a promising solution where the network is able to autonomously react and adapt to different scenarios with minimal human intervention and thus providing an operational cost reduction. Self-organization can be applied to tasks such as (*i*) network planning so that the network automatically decides the network/cell configuration in preoperational state, (*ii*) radio resource management to continuously self-optimize the radio interface performance during network operation, and (*iii*) network maintenance to provide automatic responses to network failures.

Therefore, SON is an ambitious concept that certainly will be present in future mobile communications networks. Nevertheless, some open challenges have to be further studied such as the tradeoff between the performance gains and the cost of implementation, the convergence of SON algorithms in a reasonable time, or the robustness of the algorithms to corrupt or badly configured inputs.

3.- *Flexible spectrum management.* Future wireless scenarios will probably exist in a different spectrum regulatory framework to the current perspective, where a certain piece of spectrum is allocated to a spectrum licensee for its exclusive usage during a certain period of time. Clearly, this model is inefficient in terms of spectrum utilization since some parts of the spectrum could remain unused during large periods of time whereas other could be heavily used (compare, for instance, the low usage of military reserved bands with mobile communications bands). Therefore, different spectrum regulatory models have been identified [6] being currently under consideration by the spectrum regulatory bodies. Briefly, the taxonomy of spectrum regulatory models identifies three models:

- The *Dynamic Exclusive Use Model* allocates the spectrum bands for exclusive use of services and operators along time and space. Primary operators (i.e., the licensees of the spectrum) have the usage right of a specific band but, within this model, they agree to trade their spectrum with other spectrum licensees (of the same or different service).

- The *Hierarchical Access Model* divides users between primary users (licensed users) and secondary opportunistic cognitive radio [7] users. This model mainly focuses on how secondary users may use the primary spectrum without interfering primary users' communications. In this sense, secondary users fill the spatial and temporal opportunities that primary usage of the spectrum generates.

- The *Commons Model* promotes an open sharing of the spectrum even without the control of governmental regulation bodies in some cases. This model would achieve the greatest spectrum access efficiency since any piece of the spectrum (licensed or not) would be shared spatially and temporally by primary and secondary users, in principle without any regulatory barrier. Within this new spectrum access paradigm, the so called *Private Commons* [6] has arisen, where primary spectrum owners agree to lease their spectrum to secondary markets in spectrum (i.e., potential opportunities of secondary usage of the spectrum) to be used in an infrastructure-less manner (e.g., opportunistic ad-hoc cognitive radio networks). The main difference with the *Hierarchical Access Model* is that, here, primary operators are keen to open their spectrum and to generate spectrum opportunities for secondary usage, since they may charge a fee for each commercial secondary spectrum access.

Certainly, all these models introduce new variables in the deployment of next 4G networks that have to be appropriately managed to achieve a full flexible spectrum usage. For instance, the way in which the spectrum is employed in the cellular network could be affected by the regulatory framework.

In all, it is expected that 4G mobile networks include very promising features addressed to improve users' experience through a faster and a better quality mobile service, and to wisely utilize scarce radio

resources in the network. However, several challenges regarding the implementation, deployment, and management of these networks are open.

One important challenge is the way in which OFDMA frequency resources are assigned to cells to achieve the highest spectral efficiency. Current strategies to manage the spectrum in cellular networks are based on laborious frequency planning tasks that assign spectrum resources to macrocells off-line during network deployment. With them, the network operator aims at covering the maximum demand at any place of the service area, and at reducing the intercell interference. Usually, different frequency reuse factors (FRF) are deployed, where the spectrum is distributed among macrocells following a fixed regular pattern. Normally, this spectrum assignment remains unaltered until new infrastructure is added to the system, when a tedious manual frequency planning is repeated. Although, frequency planning could be used in a 4G OFDMA-based radio interface, given the context before, there are mainly two reasons to think that further research is needed in the spectrum assignment tasks:

First, the spatial distribution of the traffic may be quite different within the operation area (e.g., a very high traffic load may be experimented in a business subarea in the mornings and very a low one in a residential subarea, whereas the opposite happens in the evenings). Frequency planning tasks that distribute frequency resources homogeneously between cells can be clearly inefficient in these situations, leading to lack of spectrum resources in some of the cells and underutilization of the spectrum in others. Also, under the flexible spectrum management frameworks explained previously, opportunities for secondary usage of the spectrum could not be exploited, reducing, in a global sense, the achieved spectral efficiency.

Second, frequency planning could become intractable with the advent of new cellular deployments based on femtocells, which are decentralized in nature and probably composed of a huge number of nodes located at random (user-decided) positions. Hence, a manual configuration of the spectrum assignment per femtocell as in the macrocell deployment is not practical.

This thesis aims at contributing on the research in this field as explained next.

## 1.1 Scope of the thesis

The high spectrum flexibility of an OFDMA radio interface made us to think, three years ago when this work started, that it would be possible to assign OFDMA frequency subchannels (groups of contiguous OFDMA subcarriers) to cells in a dynamic way rather than in the fixed fashion of frequency planning schemes. Thus, our thesis work has been focused on designing a practical framework for Dynamic Spectrum Assignment (DSA) that contributes to the spectrum assignment research in the context of next generation multicell OFDMA scenarios expected in future.

Particularly, we have considered several scenarios including macrocell deployments and an evolved spectrum regulatory framework based on Private Commons where primary operators concern themselves to perform Dynamic Spectrum Assignment with the objective of maximizing overall spectrum usage. For this purpose, primary operators attempt to guarantee the Quality of Service (QoS) of their users with the minimum spectrum whereas, on the other hand, attempt to release pieces of spectrum in large geographical areas to create spectrum access opportunities. In this sense, both primary and secondary users exploit spectrum opportunities at maximum, and thus, the primary operator obtains an additional revenue stream from spectrum trading activities with the secondary spectrum market. This attempt to

implement the Private Commons framework from a primary operator point of view, at least in the above mentioned OFDMA multicell environment, has not been found previously to the work and derived publications of this thesis. On the other hand, the dynamic spectrum management in a highly decentralized scenario such as those including femtocell deployments has been also considered.

Having this in mind, our DSA framework is envisaged to:

1.- adapt the frequency resources given to each cell to primary users' QoS requirements, taking into account the medium-term spatial and temporal variations of the traffic load.

2.- mitigate the intercell interference in order to increase the capacity per frequency resource (usually limited by the Signal to Interference plus Noise Ratio).

3.- pool and release pieces of unnecessary primary spectrum in a given region at a given point of the time so that they can be used by a secondary spectrum market.

4.- solve the spectrum assignment in femtocell scenarios out of the scope of classical frequency planning strategies.

Following the trend for the management of future networks, proposed DSA framework is based on self-organization. Hence, it constitutes a solution that autonomously observes the network performance and is able to detect changes in the traffic load a medium-term fashion (i.e., tens of seconds or minutes), determining the instants when current cell-by-cell spectrum assignment in use is not longer valid.

Two functional architectures of the framework have been designed. One is a centralized approach where a centralized entity performs the spectrum assignment simultaneously for a set of cells, taking advantage of handling a global perspective of the cells' status. On the other hand, a decentralized approach has been also proposed where the spectrum assignment is autonomously performed at each cell taking as input some limited information from the environment. Thus, this latter approach seems more adequate for femtocell deployments where a high degree of autonomy in the actions taken by each node is required.

The decision of the cell-by-cell spectrum assignment resides on a DSA algorithm. Two types of DSA algorithms have been developed along this work. One, based on heuristics, constitutes a simple but effective sub-optimal strategy to determine the spectrum assignment based on estimations of the capacity needed by each cell to cope with a given traffic load. On the other side, it has been recently argued [8] that learning could improve the performance of the self-organization strategies by effectively exploiting knowledge acquired in the past. As a result, we have introduced learning in the framework through a DSA algorithm based on reinforcement learning, which, at the same time, assures an optimal behavior in terms of maximization of a reward signal conveniently designed to cope with DSA framework objectives. Certainly learning has not been exploited at a large extent to self-organize the spectrum of an operator, and then, jointly with the DSA framework this constitutes another major contribution of this work.

Finally, it is worth to underline that the DSA framework focuses on the downlink of the cellular network (i.e., the link between base stations and users' terminals). However, we think that the framework can be easily extended to uplink direction, obviously taking into account some restrictions that may arise on the way that OFDMA frequency resources have to be assigned to cells (for instance, in order to ease terminals' implementation, OFDMA frequency resources assigned to a cell may require to be contiguous

in frequency). Also, constant transmission power per frequency resource has been considered (i.e., there is not power control). This has been motivated because it has been concluded only marginal improvements can be obtained with non-constant power assignment [9]. Nevertheless, the proposed framework would allow for the implementation power control if needed.

## 1.2 Outline of the thesis

This document is roughly organized into three parts. The first part is devoted to present a specific background to the work of the thesis by defining its technical foundations and reviewing previous state-of-the-art. Chapters 2 and 3 address these points respectively. Chapter 2 describes the technical foundations of the work derived from this thesis, which are OFDMA, self-organization and reinforcement learning. On the other hand, Chapter 3 examines strategies for spectrum assignment in OFDMA networks, or the usage of self-organization and reinforcement learning in that context. Part of the proposals given in that chapter will be used as reference cases for the performance comparison results of this thesis.

The second part of this document is devoted to present the technical proposals derived of the thesis work. Hence, Chapter 4 describes the proposed DSA framework, focusing on the description of the centralized and decentralized functional architectures, each one of the proposed functional entities, and practical implementation issues such as interfaces and signaling requirements. Once the DSA framework has been defined, Chapter 5 is devoted to present the proposed DSA algorithms that run on the DSA framework and decide the most appropriate cell-by-cell spectrum assignment at a given point of the time. Two methodologies have been considered as foundations of the DSA algorithms, heuristics and reinforcement learning, giving place, respectively, to the HEUR-DSA and RL-DSA algorithms explained in the chapter. Also, some initial validation results of the algorithms are given.

The third part of this dissertation presents the evaluation methodology and obtained results. Chapter 6 includes a description of the simulation tool developed and employed to execute dynamic simulations, a definition of the performance metrics used to compare different spectrum assignment strategies, and a description of the application scenarios where the performance comparison results were obtained. Then, Chapter 7 presents those results, where several strategies presented in the state-of-the-art given in Chapter 3 are compared with the centralized and decentralized versions of the DSA framework, running either HEUR-DSA or RL-DSA.

Finally, Chapter 8 summarizes final conclusions and suggests future work.

## 1.3 Publications

The following publications have been derived from this thesis work.

**Journal Publications**

[J1] F. Bernardo, R. Agustí, J. Pérez-Romero, and O. Sallent, “An Application of Reinforcement Learning for Efficient Spectrum Usage in Next Generation Mobile Cellular Networks”, Accepted on IEEE Transactions on Systems, MAN and Cybernetics—Part C: Applications and Reviews. 2010

[J2] F. Bernardo, R. Agustí, J. Pérez-Romero, and O. Sallent "Dynamic Spectrum Assignment in Multicell OFDMA Networks enabling a Secondary Spectrum Usage", Wireless Communications and Mobile Computing (WCMC) Journal, vol.9(11), pp. 1502-1519. Nov. 2009. DOI: 10.1002/wcm.760

**Conference Publications**

[C1] F. Bernardo, R. Agustí, J. Cordero, C. Crespo, "Self-optimization of Spectrum Assignment and Transmission Power in OFDMA Femtocells". Accepted on Sixth Advanced International Conference on Telecommunications (AICT 2010). Mayo 9 - 15, 2010 - Barcelona, Spain.

[C2] F. Bernardo, R. Agustí, J. Pérez-Romero, and O. Sallent "Distributed Spectrum Management based on Reinforcement Learning", 4th International Conference on Cognitive Radio Oriented Wireless Networks and Communications, CROWNCOM2009, Hannover, Jun, 22-24 2009. ISBN: 978-1-4244-3423-7. DOI: 10.1109/CROWNCOM.2009.5189161

[C3] F. Bernardo, R. Agustí, J. Pérez-Romero, and O. Sallent "A Self-organized Spectrum Assignment Strategy in Next Generation OFDMA Networks providing Secondary Spectrum Access", International Conference on Communications (ICC). Dresden, Germany, 14-18 June 2009. ISBN: 978-1-4244-3435-0. DOI: 10.1109/ICC.2009.5198667

[C4] F. Bernardo, R. Agustí, J. Pérez-Romero, and O. Sallent "Temporal and Spatial Spectrum Assignment in Next Generation OFDMA Networks through Reinforcement Learning", Vehicular Technology Conference (VTC)-2009Spring, Barcelona, Spain, 26-29 Apr. 2009. ISBN: 978-1-4244-2517-4. DOI: 10.1109/VETECS.2009.5073876

[C5] F. Bernardo, R. Agustí, J. Pérez-Romero, and O. Sallent "A Novel framework for Dynamic Spectrum Assignment in MultiCell OFDMA Networks based on Reinforcement Learning", Wireless Communications and Networking Conference 2009, Budapest, Hungary, 5-8 Apr. 2009. ISBN: 978-1-4244-2948-6. DOI: 10.1109/WCNC.2009.4917524

[C6] F. Bernardo, J. Pérez-Romero, O.Sallent, R. Agustí, "Advanced Spectrum Management in Multicell OFDMA Networks Enabling Cognitive Radio Usage", *IEEE international conference on Wireless Communications and Networking Conference (WCNC)*, pp. 1927 - 1932 Las Vegas, USA, March 2008, ISBN: 978-1-4244-1997-5. DOI: 10.1109/WCNC.2008.343

# 2

# TECHNICAL FUNDAMENTS

## Outline

# 2 Technical fundaments

This chapter reviews the technical foundations of the work derived from this thesis. Hence, the chapter presents the concepts behind the three technical pillars that sustain the proposals of this thesis: (1) Orthogonal Frequency Division Multiple Access (OFDMA) as a radio access interface for 4G networks with sufficient spectrum flexibility, (2) self-organization as the basis for the proposed spectrum assignment frameworks and, (3) Reinforcement Learning as a novel methodology to implement spectrum assignment strategies.

Accordingly, the chapter is divided into three parts. Section 2.1 describes OFDMA concepts, covering its benefits from a mobile communications system point of view and physical and link layer implementations. Also, its application to current 4G standards is reviewed. Section 2.2 examines the self-organization concept and its application to 4G systems. Concretely, the Self-organizing Network concept, as a network that incorporates autonomous and adaptive mechanisms based on self-organization is presented, as well as its current view in the research community. Finally, section 2.3 describes Reinforcement Learning as an optimization methodology that naturally learns the best actions to obtain the maximum reward or payoff from a correspondent environment. This section also surveys the algorithms proposed to solve the reinforcement learning problem, and concretely ends by detailing the formulation of the particular reinforcement learning algorithm adopted for the work in this thesis.

## 2.1 OFDMA concepts

Orthogonal Frequency Division Multiple Access (OFDMA) is a multiple access technique adopted by next generation mobile communications standards (e.g., LTE, WiMAX) for their radio access interfaces. Basically, OFDMA is an extension of Orthogonal Frequency Division Multiplexing (OFDM) technique, where users' access to the radio interface is provided by assigning different OFDM subcarriers to different users' data streams.

This section is devoted to describe the OFDMA concepts and its utility in next generation cellular networks. First, OFDMA benefits that motivate their selection for next generation radio interfaces are presented as well as some drawbacks that have to be taken into account. Then, concepts related to physical layer such as OFDM transmission, and the typical physical layer structure are given. On the other hand, details about link layer procedures such as dynamic packet scheduling and dynamic link adaptation enabling reliable multiuser communications through an OFDMA radio interface are described in the third subsection. Finally, this clause is ended by reviewing how OFDMA has been adopted in LTE and WiMAX standards.

### 2.1.1 OFDMA benefits

The usage of OFDMA entails the following advantages for a mobile communication system:

- Robustness against multipath propagation: OFDMA is very robust against the intersymbolic interference produced by multipath propagation thanks to the application of the *cyclic prefix* explained in section 2.1.2.1. Moreover, signal distortion can be easily combated by means of

equalization techniques in the frequency domain, which are more efficient and less complex than the classic equalization techniques in the temporal domain, especially for broadband channels (i.e., LTE could arrive to transmission bands of 20 MHz).

- High degree of use of the assigned band: Thanks to the use of OFDM modulation (section 2.1.2.1), an OFDMA radio interface can perform transmission with a minimum spacing between the different subcarriers without affecting the recovery of the transmitted signal. On the contrary, other transmission techniques different from OFDM would require a certain guard band between the different subchannels, which would consequently reduce the degree of use of the assigned band and the spectral efficiency.

- Ease of implementation, by simple digital processing through IFFT/FFT (Inverse Fast Fourier Transform/ Fast Fourier Transform) operation (section 2.1.2.1).

- High granularity in the assignable resources: An OFDMA interface provides a high granularity since the radio interface is composed of several narrow band subchannels which could be dynamically assigned to different users in a flexible way (section 2.1.2.2). On the contrary, in other techniques like CDMA (Code Division Multiple Access), the variation in the amount of assignable resources to each user is obtained mainly by means of the modification of the spreading factor, which used to have a smaller granularity since it usually varies in powers of 2.

- Frequency diversity: In an OFDMA interface, it is possible to assign to a same user sufficiently separated OFDM subcarriers so that the state of the channel is independent. This provides frequency diversity in the transmissions of a given user and robustness against frequency selective channels.

- Multiuser diversity: In an OFDMA radio interface, the allocation of subcarriers to users is carried out dynamically, being able to change in short periods of time through scheduling strategies (section 2.1.3.1). Considering that the radio channel will show random fading in the different subcarriers, and that that fading will be independent for each user, it is possible to select, for each subcarrier and at any time instant, a user who presents a good state of the channel (e.g., that perceives an acceptable signal to noise ratio). This phenomenon is known as *multiuser diversity* and translates in a better usage of the available band to obtain a greater transmission data rate, especially for high fading channels and a high number of users.

- Flexibility in the assigned band: OFDMA provides a simple way to accommodate different transmission speeds for the different users based on their requirements, by changing the amount of subcarriers assigned to each user.

On the other hand, OFDMA presents some drawbacks:

- High Peak to Average Power Ratio (PAPR): One of the disadvantages of the OFDMA techniques is that the transmitted instantaneous power can be significantly superior to the average power, which creates problems for the linearity of the power amplifiers, limiting their efficiency and increasing their cost. This is one of the reasons why OFDMA has not been adopted for the uplink of LTE system, since building high linear power amplifiers for the terminals would be very

costly. Instead, the so called Single Carrier Frequency Division Multiple Access (SC-FDMA) was adopted [10].

- Sensitivity to frequency errors: In case of frequency shifts of OFDM subcarriers with respect to their reference frequency, a certain loss of orthogonality between subcarriers is produced, which translates into interference between subcarriers. The reasons for these errors in frequency can be diverse such as, for instance, the stability of the oscillators, the Doppler shift associated to the mobility of the terminals, etc. For this reason, synchronization takes an important role in OFDM based systems.

## 2.1.2 Physical layer

### 2.1.2.1 OFDM modulation

Orthogonal Frequency Division Multiplexing (OFDM) is a well-known frequency division multiplexing technique in literature [11]. The basic principle of OFDM is the conversion of a high-rate serial data stream to multiple low-rate substreams transmitted over different subcarriers. As described in this section, OFDM features are: (*i*) the usage of multiple narrowband subcarriers to carry the information, (*ii*) the orthogonality between subcarriers, which assures that the information sent over a given subcarrier does not affect the others, and (*iii*) the introduction of guard times to assure orthogonality, combat channel delay spread and reduce the complexity of the receptor. These features bring important benefits such as, among others, frequency diversity, multipath propagation robustness and high spectral usage that are generally acknowledged to be well suited the mobile radio environment, especially for broadband systems.

The main parameters of OFDM modulation are:

1.- The OFDM symbol duration $T_s$.

2.- The subcarrier separation $\Delta f$, which by definition is $\Delta f \triangleq 1/T_s$.

3.- The number of subcarriers $N$. Taking into account the sub-carrier separation, this yields to a total bandwidth of the OFDM signal of $BW = N\Delta f$.

4.- The conforming pulse of the modulation used in each subcarrier, which, classically consists of a time rectangular pulse.

As shown later in this subsection, the OFDM signal is obtained by adding the contributions of $K$ subcarriers. Mathematically, each subcarrier is formulated in baseband as a complex signal in time domain as:

$$x_n(t) = e^{j2\pi n\Delta f t} rect_{T_S}(t) \qquad 0 \le n \le N-1, \tag{2.1}$$

where $f_n = n\Delta f$ is the frequency of the subcarrier $n$. Finally, $rect_{T_S}(t)$ stands for a pulse of rectangular shape between 0 and $T_s$. Notice that two subcarriers $m$ and $n$ are orthogonal in a temporal interval of duration $T_s$ since the integration of their product is null except when $m = n$ as follows:

$$R_{x_m,x_k}(t)=\frac{1}{T_S}\int_0^{T_S}x_m(t)\,x_n^*(t)\,dt=\frac{1}{T_S}\int_0^{T_S}e^{j2\pi(m-n)\Delta ft}dt=\frac{1}{T_S}\int_0^{T_S}e^{j2\pi(m-n)\frac{t}{T_S}}dt=\begin{cases}1 \text{ if } m=n.\\ 0 \text{ if } m\neq n.\end{cases} \quad (2.2)$$

Figure 2.1 shows the representation of an OFDM signal in time and frequency domain. Note that in the frequency domain, only one of the subcarriers contributes to the OFDM signal for each frequency multiple of $1/T_s$, whereas the rest of subcarriers are null at that frequency. This is because the 'sinc' shape obtained in frequency domain for a given subcarrier has zeros matching with the frequency allocation of the other subcarriers, thanks to the inverse relation between subcarrier spacing and OFDM symbol duration.

Another important feature of OFDM modulation is the introduction of guard times. Usually it is added at the beginning of the OFDM symbol and carries a copy of the final OFDM symbol information. For this reason it is commonly referred as *cyclic prefix*. This prefix helps to reduce the Inter-Carrier Interference (ICI) and thus helps to assure orthogonality in frequency domain. Moreover, besides the reduction of the bit rate per sub-carriers, cyclic prefix helps to reduce the Inter-Symbol Interference (ISI) in time domain if it is wisely selected according to channel characteristics.

The more important channel characterization parameters are the delay spread ($\tau$) and the Doppler shift ($f_d$). The former reflects the dispersion of the energy of the symbols due to multi-path channel. The latter affects the coherence time of the channel ($1/f_d$), which reflects the time the channel can be considered invariant. In order to avoid ISI, the duration of the cyclic prefix, $T_g$, must be selected to be greater than the delay spread. In this case, the entire energy dispersion is absorbed in the guard time. Moreover, the OFDM symbol duration should be much greater than delay spread but much lower than the coherence time in order to consider the channel constant during OFDM symbol duration. In the same way, due to the inverse relation between subcarrier separation and OFDM symbol duration, similar considerations can be made for the subcarrier separation, where $\Delta f$ must be much lower than the

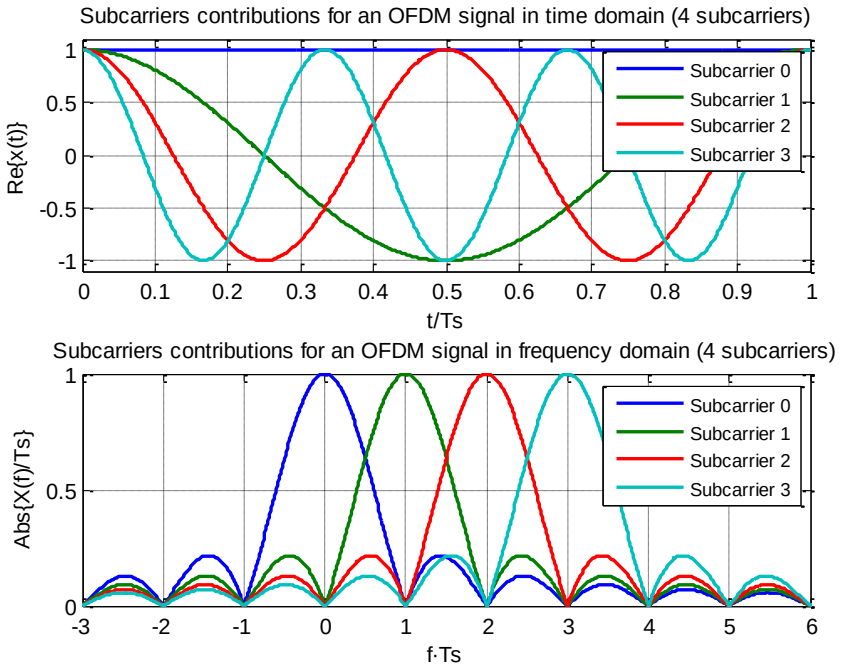


**Figure 2.1 OFDM subcarriers in time and frequency domain**

coherence bandwidth ($1/\tau$) in order to assure that each subcarrier can be modeled as 'flat' response channel. All this is expressed in the following expressions:

$$T_g > \tau \tag{2.3}$$

$$\tau << T << \frac{1}{f_d} \tag{2.4}$$

$$f_d << \Delta f << \frac{1}{\tau} \tag{2.5}$$

These relationships have two major implications. On the one hand, since subcarriers are independent to each other, different linear modulation schemes can be applied to each subcarrier (or group of them) and in consequence achieve the best spectrum efficiency in each case. On the other hand, thanks to orthogonality and guard times, each subcarrier can be considered as a flat narrowband channel free of fast fading impairments (delay spread).

Finally, regarding OFDM implementation, OFDM signal generation can be made by adding the output of $N$ parallel modulators each one over a different subcarrier. However, practical implementations use $N$-point IFFT/FFT (Inverse Fast Fourier Transform/ Fast Fourier Transform) operation which allows easy implementation by processing the digital base band OFDM signal. IFFT is used in transmitter while FFT operation is used in reception. Figure 2.2 shows the whole process.

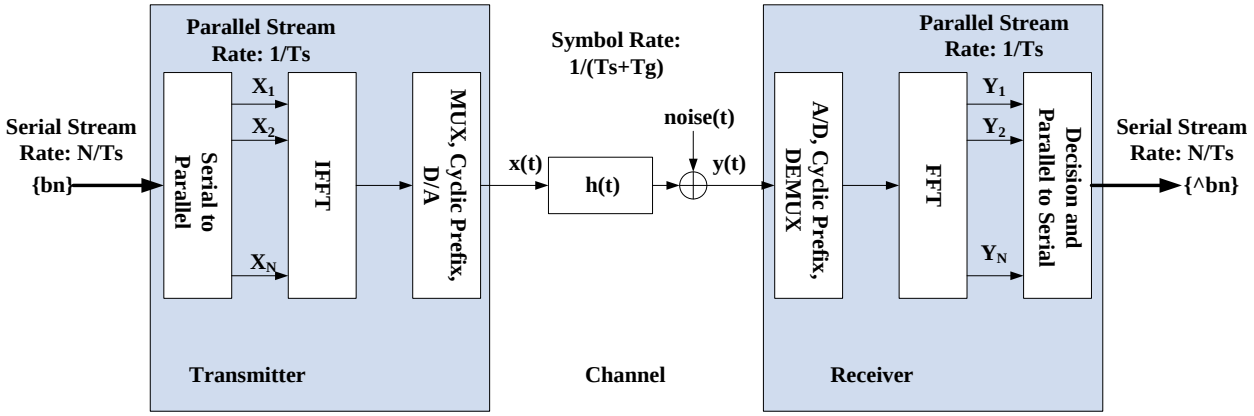


**Figure 2.2 Baseband model for OFDM transmitter and receiver implementation**

Due to cyclic prefix and orthogonality, each sub-carrier is affected by a narrowband flat channel. Thus it can be shown [12] that

$$Y_n = X_n H_n + noise_n, \tag{2.6}$$

where each point of the IFFT/FFT is affected by the channel response in its frequency. This simplifies very much the receptor implementation, as equalization process is simple and independent for each sub-carrier.

### 2.1.2.2 Physical layer structure

Orthogonal Frequency Division Multiplexing Access (OFDMA) exploits OFDM modulation as a multi-user access technology. Typically, the radio resources in OFDMA radio interface are divided in both time and frequency building a time-frequency grid as shown in Figure 2.3, where the minimum radio

resource that can be assigned to a user is usually named as a Resource Block (RB). In frequency, the whole available bandwidth is divided into groups of adjacent subcarriers or *chunks* whereas in time it is divided into *frames*.

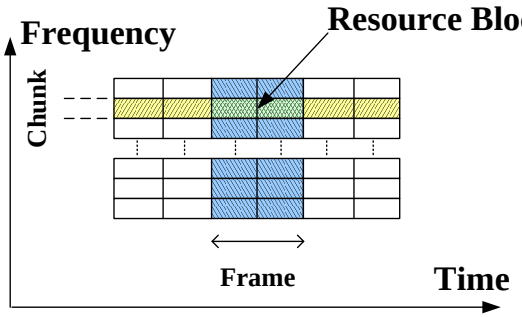


**Figure 2.3 Time-Frequency grid in OFDMA radio interfaces**

OFDMA allows for partitioning the time and frequency resources into the OFDM symbols, and the OFDM subcarriers (or chunks), in a way that a user can be assigned a frame to transmit over a subcarrier (or chunk). Figure 2.4 illustrates this concept, where $U$ users are easily multiplexed into the OFDM signal by inserting the $N_u$ symbols corresponding to the $u$-th user into $N_u$ OFDM subcarriers (i.e., $N_u$ points of the IFFT). Analogously, each user terminal only takes the information carried in the subcarriers assigned to it (notice that some signaling is needed to notify which subcarriers are devoted to each user)[1]. Finally, it is worth to remark that subcarriers assigned to a user do not need to be contiguous as depicted in the figure for simplicity.

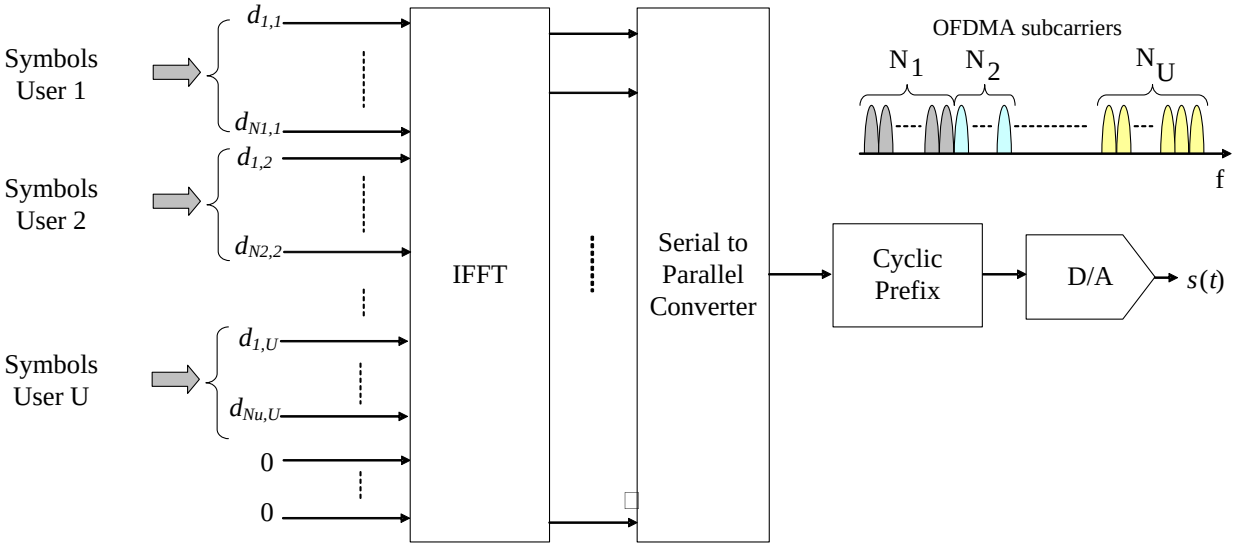


**Figure 2.4 User multiplexing in OFDMA**

[1] In this thesis, it is considered that the minimum assignable block in frequency is a chunk, and thus user multiplexing would be performed with a chunk granularity.

## 2.1.3 Link layer

In the following two important radio resource management functionalities related with OFDMA link layer are described. They are dynamic packet scheduling and dynamic link adaptation. The former is devoted to decide how users are multiplexed into the whole OFDM signal. The latter address the selection of the appropriate modulation and coding schemes for each user so that variable data rate with reliable transmission is attained.

### 2.1.3.1 Dynamic packet scheduling

Dynamic packet scheduling in OFDMA consists of dynamically assigning radio resources to users so that they can orderly make their transmissions through the radio interface. Packet scheduling entails several subtasks like deciding, which users are going to transmit, in which resources of the radio interface, and with which modulation and codification scheme. This last task is known as dynamic link adaptation and will be separately described in section 2.1.3.2. To perform abovementioned tasks, packet scheduling considers, among others, the Quality of Service (QoS) requirements of associated to the users' services, the channel quality for each terminal, or the state of buffers.

The main benefit of dynamic packet scheduling is that it facilitates the efficiently exploitation of the radio resources in an OFDMA interface. On the one hand, in the temporal domain, it is possible to exploit the so called multiuser diversity, where, due to the existence of several users, it is more probable that some of them experience a good quality of the channel in a given frame. Thus, if the packet scheduler considers the channel status of each user (known as *channel-aware scheduling*), it could, for instance, assign the channel to the user with the best channel quality and then it would be possible to take advantage at any time of the best channel quality. This multiuser diversity improvement is greater as greater is the number of users and faster are the variations of the channel.

On the other hand, it is possible to extend the multiuser diversity concept to the frequency domain Figure 2.5, where the variations of the channel due to frequency selective fading can also be considered. This is particularly appreciable in cellular systems where the coherence bandwidth is typically smaller than the bandwidth of the system[2]. Figure depicts the channel for two users. Note that channel response is different for each user in both frequency and time the channel, so that if the user to transmit is properly selected, then it is possible for the system to work more efficiently (e.g., next to its capacity).

Usually packet scheduling function is placed at the base station in 4G next generation systems. The reason is that in this way it is possible for the scheduler to operate in a short-term scale (around milliseconds) allowing low latency response and fast adaptation to channel variations. This centralized scheme also assures that transmissions for users connected to the same base station do not interfere each other (i.e. there is orthogonality between them). However, there is a cost in signaling between users and the base station to report channel quality, buffers occupancies, etc. It is important to remark that, although the scheduling functionality is executed at each base station, it could consider some restrictions or preferences when assigning the chunks to users. For instance, in case of LTE, the functionality Intercell Interference Coordination (ICIC) [14] provides some coordination between schedulers in different cells with the aim of coordinating their respective decisions and mitigate intercell interference. This also opens

[2] Channel model Typical Urban (TU table C.3.3 on [13]) has a coherence bandwidth of approximately 190 kHz, whereas 4G systems will be deployed with system bandwidth above 1 MHz (e.g., minimum deployment bandwidth in LTE is 1.25 MHz)

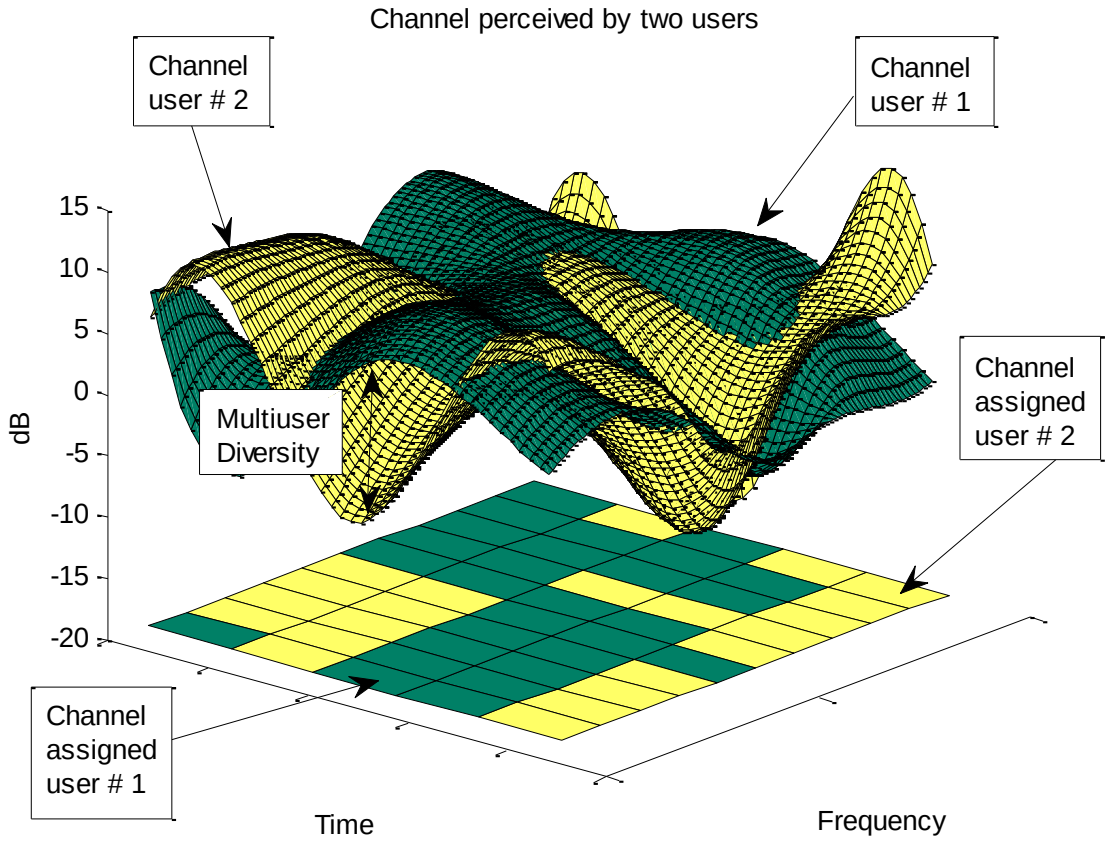


**Figure 2.5 Illustration of the multiuser diversity and its possible exploitation through scheduling.**

the way for spectrum assignment framework and strategies such as the proposed in this thesis in Chapters 4 and 5.

### 2.1.3.2 Dynamic link adaptation

One of the tasks of the packet scheduler is the selection of the most suitable modulation and codification format at any moment for the transmissions of each user. This decision process is known as *dynamic link adaptation*. Usually, a mechanism based on the control of the binary rate is implemented (*rate control*), that consists of varying the assigned binary rate to users based on their estimated channel quality. In consequence, the user will perceive high a binary rate if the channel quality is good, and low rate if the channel quality is bad.

Basically, this procedure is generally depicted in Figure 2.6. Dynamic link adaptation is based on the estimation of the channel quality, which, for the UL, is obtained from the reception of predefined reference signals and, for the DL, is given in the information contained in Channel Status Reports sent by users. From this estimation it is possible to determine a modulation and coding scheme to guarantee a certain Bit Rate Error (BER) that it will depend on the offered service and available coding schemes. As an example, Figure 2.7 shows a generic table to obtain a BER of $10^{-2}$ when attaining the maximum transmission speed for each value of Signal to Interference plus Noise Ratio (SINR) (considering that the estimation of the channel quality could be translated in different values of SINR). Also, the Shannon theoretical capacity limit is depicted. It can be seen in figure that robust modulation and coding schemes like QPSK are uses for low SINRs, at the cost of having a low spectral efficiency and therefore reducing the binary rate.

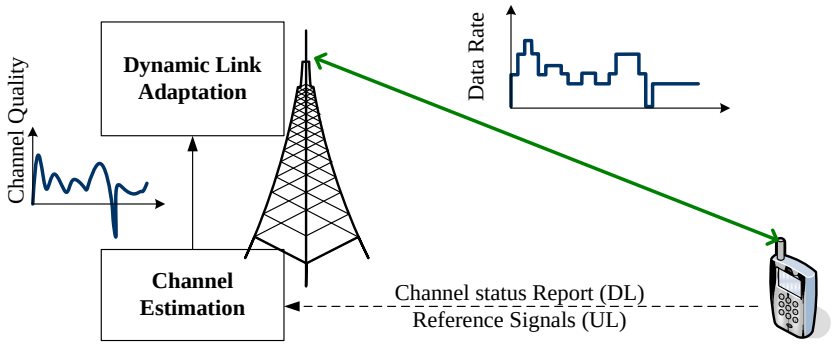


**Figure 2.6 Dynamic Link Adaptation based on rate control**

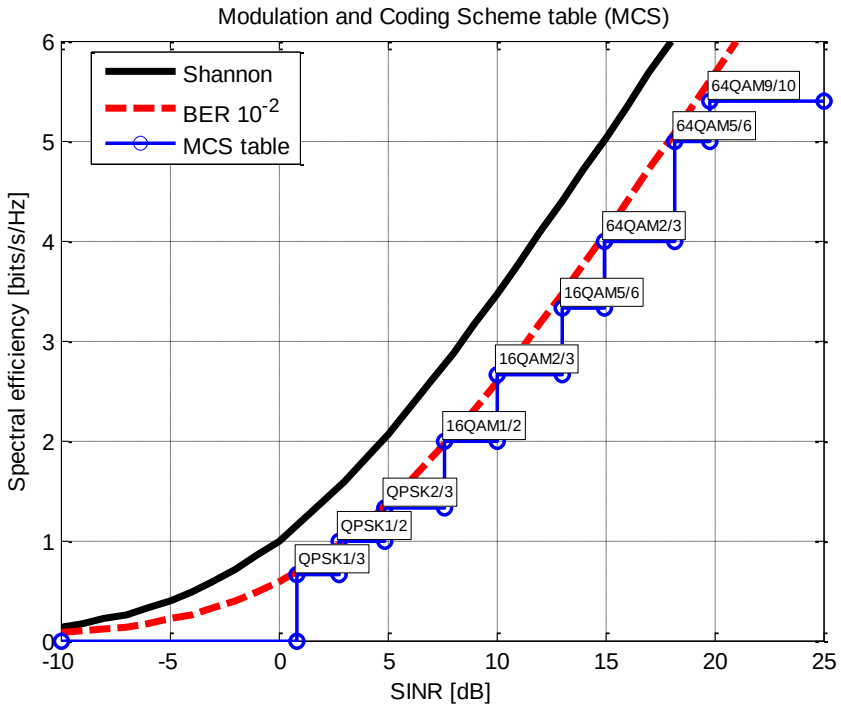


**Figure 2.7 Example of a Modulation and Coding Scheme table**

## 2.1.4 Standardized 4G radio access interfaces based on OFDMA

Some current 4G systems use OFDMA as multiplexing technology in the radio interface. The most representative worldwide are 3GPP LTE [14] and IEEE802.16 WiMAX [15], whose OFDMA implementations are detailed next.

### 2.1.4.1 OFDMA implementation in LTE

3GPP LTE devises the usage of OFDMA technology for the downlink direction of the radio interface. However, in uplink OFDMA was dismissed due to its intrinsic high Peak to Average Power Ratio (PAPR) what would increase the cost of the terminal's power amplifiers. Instead, the Single Carrier-Frequency Division Multiple Access (SC-FDMA) was adopted [14]. Nevertheless, SC-FDMA operation is similar to OFDMA where a Discrete Fourier Transform (DFT) pre-codification and IFFT point shifting process is previously introduced to normal OFDMA transmission. As a result, SC-FDMA restricts the chunks that can be assigned to a user to be contiguous in frequency.

Focusing then in LTE downlink implementation, the subcarrier spacing is $\Delta f$ =15 kHz. Subcarriers are grouped in blocks of 12 contiguous subcarriers (i.e. chunks) with a resultant bandwidth of 180 kHz. This constitutes the minimum assignable frequency resource in LTE to a user. The flexibility in the occupied bandwidth is given by the number of blocks used $N_B$. Concretely, $N_B$ could takes the values of {6, 15, 25, 50, 75, 100} blocks, corresponding to LTE deployment bandwidths of {1.4, 3, 5, 10, 15, 20} MHz.

In terms of IFFT/FFT implementation, for each occupied bandwidth, the number of points is given by $N$ ={128, 256, 512, 1024, 1536, 2048} and corresponding sampling frequency of $f_m$ ={1.92, 3.84, 7.68, 15.36, 23.04, 30.72} MHz. These values were selected to ease the compatibility between LTE and 3GPP UMTS system. On the other hand, OFDM symbol duration is $T_s = 1/\Delta f$ =66.67 μs and cyclic prefix could take values in the set {5.21, 4.68, 16.7} μs.

Finally, modulations envisaged for LTE downlink are QPSK, 16 QAM and 64 QAM, corresponding to 2, 4 and 6 bits per symbol respectively.

#### 2.1.4.2 OFDMA implementation on WiMAX

WiMAX standard also makes use of OFDMA as radio access technology. The fixed and mobile versions of WiMAX have slightly different implementations of the OFDM physical layer. Fixed WiMAX, which is based on IEEE 802.16-2004 [16], uses a 256 FFT-based OFDM physical layer. Mobile WiMAX, which is based on the IEEE 802.16e-2005 standard [17], uses a scalable OFDMA-based physical layer. In the case of mobile WiMAX, the FFT sizes can vary from 128 bits to 2048 bits. The available subcarriers may be divided into several groups of subcarriers called subchannels.

Fixed WiMAX based allows a limited form of subchannelization in the uplink only. The standard defines 16 subchannels, where 1, 2, 4, 8, or all sets can be assigned to a subscriber station (SS) in the uplink. Uplink subchannelization in fixed WiMAX allows subscriber stations to transmit using only a fraction of the bandwidth allocated to it by the base station, which provides link budget improvements that can be used to enhance range performance and/or improve battery life of subscriber stations.

Mobile WiMAX, however, allows subchannelization in both the uplink and the downlink, and here, subchannels form the minimum frequency resource unit allocated by the base station. Therefore, different subchannels may be allocated to different users as a multiple-access mechanism following then an OFDMA access strategy. Subchannels may be constituted using either contiguous subcarriers or subcarriers pseudo-randomly distributed across the frequency spectrum. Subchannels formed using distributed subcarriers provide more frequency diversity, which is particularly useful for mobile applications.

WiMAX defines several subchannelization schemes based on distributed carriers for both the uplink and the downlink. One, called partial usage of subcarriers (PUSC) is mandatory for all mobile WiMAX implementations. The initial WiMAX profiles define 15 and 17 subchannels for the downlink and the uplink, respectively, for PUSC operation in 5 MHz bandwidth. For 10 MHz operation, it is 30 and 35 channels, respectively. The subchannelization scheme based on contiguous subcarriers in WiMAX is called band adaptive modulation and coding (AMC). Although frequency diversity is lost, band AMC allows system designers to exploit multiuser diversity, allocating subchannels to users based on their frequency response. In general, contiguous subchannels are more suited for fixed and low-mobility applications.

Most WiMAX deployments are likely to be in TDD (Time Division Duplex) mode because of its advantages. TDD allows for a more flexible sharing of bandwidth between uplink and downlink, does not require paired spectrum, has a reciprocal channel that can be exploited for spatial processing, and has a simpler transceiver design. The downside of TDD is the need for synchronization across multiple base stations to ensure interference-free coexistence. Paired band regulations, however, may force some operators to deploy WiMAX in FDD (Frequency Division Duplex) mode.

## 2.2 Self-organization concepts

Self-organization is the ability of a system composed of several entities to adopt a particular structure and perform certain functions to fulfill a global purpose without any external supervisor or central dedicated control entity [18]. Intuitive examples of self-organization are swarms of ants looking for food, or schools of fish protecting against predators.

The main characteristics of a self-organized system are its *distributed* nature and the *localized* interactivity between the elements of the system. That is, each entity performs its operation based only on the information retrieved from other entities in its vicinity. Hence, overall system's organization and performance is achieved from an *autonomous* behavior of each entity that, from the *experience* acquired from a variable *environment*, decides the proper *actions* to adapt to it.

A classical benefit of self-organized systems is its *robustness* against failures, since the operation of the system could maintain even in the case of damage of one of the entities. Other benefit is its degree of *scalability*, where other entities can be added without compromising the performance of the entire system, especially if the number of entities becomes very large.

In the field of radio access networks, self-organization can be applied to network planning, deployment, optimization, and maintenance [19]. Design paradigms [20] and several use cases [21] have been identified with regard to typical macrocell cellular networks. In this context, self-organization is regarded in a more general level, where it is expected that network procedures are executed automatically without human intervention, and with convenient information retrieved from the network status. The main objective is to achieve operational and capital expenditures reductions. Besides, distributed execution is not always required.

Radio access networks that include self-organization principles are commonly known as Self-Organizing Networks (SON) and their main concepts are described in section 2.2.1. Finally, this clause ends in section 2.2.2 by reviewing certain activities of several projects and standardization bodies steered to study the automation of cellular mobile network procedures.

### 2.2.1 Self-Organizing Networks (SON)

A Self-organizing Network (SON) is a communication network which supports self-x functionalities, e.g. self-configuration or self-optimization [22]. Self-x enables the automation of operational tasks, and thus it minimizes human intervention. Generally self-x functionalities are based on a loop known as self-x cycle (Figure 2.8), where input data is gathered through network monitoring and current status, then these data is processes to derive optimized parameterization, and eventually the new configuration is set up in the network.

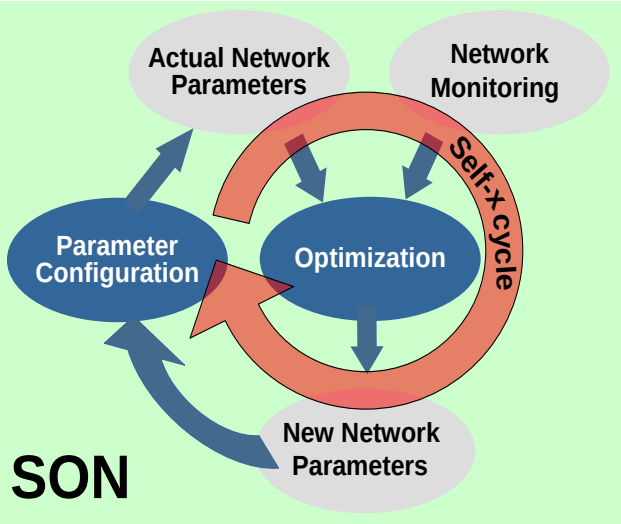


**Figure 2.8 Self-x cycle in a Self-Organizing Network**

SON is considered along its different manifestations addressed as self-x:

- Self-configuration: Self-configuration process is defined as the process where newly deployed nodes are configured by automatic installation procedures to get the necessary basic configuration for system operation. This process is performed in pre-operational state. The pre-operational state is defined as the state where the RF interface is not commercially active.

- Self-planning: Self-planning comprises the processes were radio planning parameters are assigned to a newly deployed network node. Parameters in scope of self-planning are (1) Neighbor cell relations; (2) Max TX power values of UE and eNB; (3) HO parameters, Hysteresis, trigger levels, etc..

- Self-Optimization: Self-optimization is defined as the process where UE and base station measurements and performance measurements are used to auto-tune the network. The tuning actions could mean changing parameters, thresholds, neighborhood relationships, etc. This process is accomplished in the operational state. The operational state is defined as the state where the radio interface is commercially active.

- Self-managing: Self-managing is the automation of Operation and Maintenance (OAM) tasks and workflows, i.e. shifting them from human operators to the mobile networks and their Network Elements. The burden of management would rest on a mobile network itself while human operators would only have to provide high level guidance to OAM. The network and its elements can automatically take actions based on the available information and the knowledge about what is happening in the environment, while policies and objectives govern the network OAM system.

- Self-healing: Self-healing is a SON functionality which detects problems itself and solves or mitigates these problems to avoid user impact and to significantly reduce maintenance costs. For each detected fault, appropriate alarms shall be generated by the faulty network entity. The Self-healing functionality monitors the alarms, and when it finds alarm/s which could be solved automatically, it gathers more necessary correlated information (e.g. measurements, testing results, etc) and does deep analysis, and then according to the analysis result, it triggers appropriate recovery actions to solve the fault automatically.

Therefore, self-x effectuates the improvement of the usability of future wireless access solutions ("plug&play"), and accelerates the introduction and deployment of new wireless services. In addition, self-organizing approaches may contribute to further increasing spectral efficiency, since they can be used to allocate capacity where it is needed, in contrast to today's networks, whose design is made to handle the maximum demand expected at any place in the covered area. Finally, self-x approaches target also on improvement of Quality of Service (QoS) perceived by the user. Besides the increase of spectral efficiency, this is also the optimization of interference and coverage in critical reception conditions. Hence, the main gains of self-x are expected first of all in operational expenditure (OPEX) reductions and secondly in network performance improvements.

### 2.2.2 Self-organization view in projects and standardization

The current framework of a SON is governed by the activities deployed under several research initiatives which try to develop the appropriate techniques for the automation of management and optimization tasks in the radio access network.

- The SOCRATES (Self-Optimization and self-ConfiguRATion in wirelEss networkS) project [23] is supported by the European Union under the 7th Framework Program. The general objective of SOCRATES is to develop self-organization methods in order to optimize network capacity, coverage and service quality while achieving significant OPEX (and possibly CAPEX) reductions. Although the developed solutions are likely to be more broadly applicable (e.g. to WiMAX networks), the project primarily concentrates on 3GPP's LTE radio interface (E-UTRAN).

- The E3 (End-to-End Efficiency) project [24] is supported by the European Union under the 7th Framework Program. The project is evolving current and future heterogeneous wireless system infrastructures into an integrated, scalable, efficiently managed Beyond 3G (B3G) cognitive system framework. E3 brings together European key players in the domain of cognitive radios and networks, self-organization and end-to-end reconfigurability. The key objective of the E3 project is to design, develop, prototype and showcase solutions for optimized usage of existing and future radio access resources. In particular, more flexible use of frequency spectrum, terminals, base stations and networks is addressed.

- In 3GPP Release 8 many of the signaling interfaces between network elements are standardized (open) interfaces. Significant examples in the context of SON are the X2 interface between eNBs and the S1 interface between eNB and the EPC (e.g. MME, SGW). On the other hand, for 3GPP Release 8 it has been decided that SON algorithms themselves will not be standardized. Several working groups propose contributions on this subject, that are discussed and finally relevant conclusions are extracted. In [14], the scope of self-configuring and self-optimizing functionality is defined. Both processes are described as well as the functions handled by them

## 2.3 Reinforcement Learning concepts

Reinforcement Learning (RL) comes from the field of artificial intelligence and machine learning. RL is *learning* what to do in order to maximize a numerical *reward* given that there is an interaction with an *environment* [25]. RL learns from that interaction and not from correct examples or patterns given by an

external supervisor as occurs in supervised learning (the kind of learning studied in most current research in machine learning, statistical pattern recognition, and artificial neural networks).

Reinforcement learning is not defined by characterizing learning algorithms, but by characterizing a learning problem. RL considers the whole problem of a *goal-directed* agent interacting with an uncertain environment. Thus, there are two main elements in a RL problem: the RL *agent*, the *environment* that mutually interact (Figure 2.9). All reinforcement learning agents can sense aspects or *states* of their environments, and can choose *actions* to influence their environments in order to receive *rewards* from the environment (i.e., the goal achievement). This scheme is repeated several times or steps.

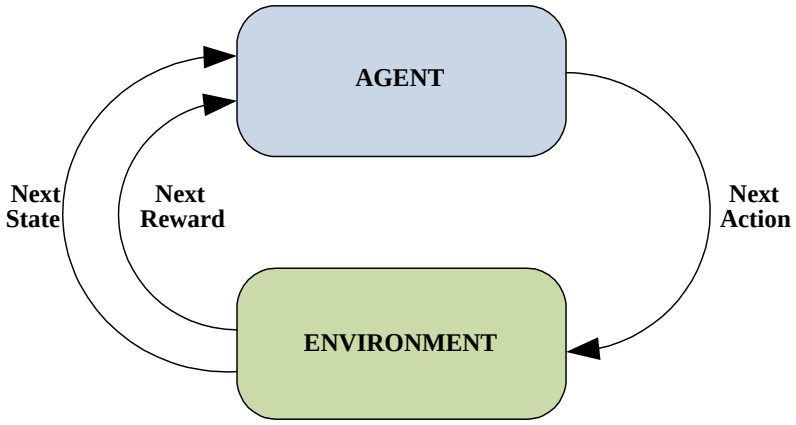


**Figure 2.9 Reinforcement Learning framework**

The *state* of the environment represents whatever information that is available to the agent from the environment. Then, it represents some measurements, data or high representation signals of the environment (e.g., in a radio environment it could represent the total inter-cell interference level, the number of frequencies assigned or discrete states that indicate whether the QoS of the users is satisfied or not).

The *reward* is a numerical representation of the goal achievement in a specific step and varies from one step to another. In fact, the agent's goal is to maximize the total amount of reward it receives in the long-run. More precisely, the agent tries to maximize the *expected return* of its actions. The expected return is a combination of all the rewards obtained for present and future actions that can be a simple summation or similar of the rewards. Then actions followed by large rewards in a specific step should be made more likely to recur, whereas actions followed by small rewards should be made less likely to recur. But it is important to set the *reference reward* in order to determine what is a large or a small reward. Typically, reference reward is an estimation of the average reward. Usually, algorithms that only take into account the last reward are said to implement *immediate reinforcement* whereas those that take into account the value are called *delayed reinforcement* algorithms.

Besides the agent and the environment there are other secondary elements of a RL system. They are a *policy*, a *reward function*, a *value function*, and, optionally, a *model of the environment*. A *policy* defines the agent's learning behavior at a given time. It is a mapping from perceived states of the environment to actions to be taken when in those states. A *reward function* defines the goal in a reinforcement learning problem by giving the numerical value of the reward for a given action and state. A *value function* specifies what is good in the long-term. The *value* of a state is the total amount of reward an agent can expect to accumulate over the future starting from that state. Finally, a *model* of the environment is something that mimics the behavior of the environment.

One of the characteristic features of RL is the trade-off between *exploration* and *exploitation*. The RL agents (those that run RL algorithms) try from interaction to discover (learn) the best actions that better solve a problem. In their execution, they will choose the best actions that gave more reward (they exploit the learning). However, the agents sometimes should select actions that are not considered the best (in this case they explore) in order to discover new actions that may lead to better solutions in future. Obviously, exploration and exploitation go in different directions in the consecution of the best reward in the short-term, but RL wisely joins both features to build a robust self-adaptive methodology that learns from environment and reach great reward in the long-term.

There are several methods to balance the exploration and the exploitation. *Action-Value* methods are those that select the actions depending on its estimated value[3]. Within these set of methods we have the greedy method that simply selects the action with the greatest value (i.e., it never explores) and the *ε-greedy* method, whose rules select (explore) from time to time a random action uniformly with a smell probability $\varepsilon$. The ε-greedy method outperforms the greedy method since the last one could get stack in sub-optimal solutions that the ε-greedy may overcome thanks to its exploratory behavior. There are other methods such as the *Softmax Action* selection that grades the probability of selecting actions from best to worse instead of exploring them uniformly as the ε-greedy method. The *reinforcement comparison* methods also grade the actions with preferences of selection but update these preferences proportionally to the difference between the actual reward and the reference or average reward[4]. Finally, *pursuit* methods weight the actions in a way that keeps the agent selecting actions towards the greedy (best) action.

### 2.3.1 Formulation of the value function

Almost all reinforcement learning algorithms are based on estimating *value functions*. These functions represent *how good* it is for the agent to be in a given state. Then they are useful to determine which are the best policies. That is, those that best learn from the environment and achieve the best reward.

As it has been said, a policy represents a mapping from the set of states $S$ and the set of actions $A$ to the probability that a specific action $a$ is taken when in state $s$ (denoted as $\pi(s,a)$). The value $V^{\pi}(s)$ of a state $s$ under a policy $\pi$ is the expected return $R_t$ when starting in $s$ and following to any other state. Formally is defined as:

$$V^{\pi}(s) = E\{R_t \mid s_t = s\}. \tag{2.7}$$

Similarly, we define the value $Q^{\pi}(s,a)$ of taking action $a$ in state $s$ under a policy $\pi$ as the expected return starting from $s$, taking the action $a$, going to another state following policy $\pi$.

[3] One common way of estimating the action value is the *simple-average* method that averages the rewards obtained for that action when it was selected. If $Q_k(a)$ denotes the average value of action $a$ after being selected $k$ times and $r_k$ the actual reward then $Q_{k+1}(a) = Q_k(a) + [1/(k+1)][r_{k+1} - Q_k(a)]$ where the initial value $Q_0$ plays an important role in the exploratory behaviour of the agent. Optimistic initial values stimulate the agent to perform exhaustive exploration at a cost of higher convergence times.

[4] Most of the RL methods use this "error estimate" philosophy where the new estimate is computed as *New_Estimate = Old_Estimate + Step_Size [Target – Old_Estimate]*.Step_Size plays a crucial role in the convergence of the methods. It is good that it decreases with time is the problem is stationary while it is better to keep it constant in non-stationary problems (usually).

$$Q^{\pi}(s,a) = E\{R_t \mid s_t = s, a_t = a\}. \tag{2.8}$$

A fundamental property of value functions used throughout RL and dynamic programming is that they satisfy particular recursive relationships. Thus they can be expressed in terms of the value of its successor states (then some prediction of the expected value can be done). This is expressed in the *Bellman equation*:

$$V^{\pi}(s) = \sum_{a} \pi(s,a) \sum_{s'} P^{a}_{ss'}[R^{a}_{ss'} + \gamma V^{\pi}(s')], \tag{2.9}$$

where $s'$ represents the next possible states and $\gamma$ a discount rate. Moreover $P^{a}_{ss'}$ is the transition probability from a state $s$ to another $s'$ given that action $a$ was selected and $R^{a}_{ss'}$ is the reward obtained with that transaction. Concretely, at a given step *t* they are defined as

$$P^{a}_{ss'} = \Pr\{s_{t+1} = s' \mid s_t = s, a_t = a\}; \tag{2.10}$$

$$R^{a}_{ss'} = E\{r_{t+1} \mid s_t = s, a_t = a, s_{t+1} = s'\}, \tag{2.11}$$

where $r_t$ is the current reward and $E$ is the expected value operator.

The *value function* is the unique solution to its Bellman equation and can be used to compute, approximate and learn the best policy $\pi^*$ that returns the largest values for each state.

## 2.3.2 Solving methods

There are three fundamental classes of methods for solving the RL problem based on the estimate of the value functions above: *dynamic programming* (DP), *Monte Carlo* methods, and *temporal-difference learning* (TD). Each one of the three methods has its strengths and weaknesses. Dynamic programming (DP) methods are very well developed mathematically, but require a complete and accurate model of the environment such as a *Markov decision process* model[5]. Monte Carlo methods do not require a model of the environment and are very simple conceptually, but are not well suited for step-by-step incremental computation. Finally, TD methods do not require an environment model and are fully incremental, but are more complex to analyze. Nevertheless TD methods joint the benefits of both DP and Monte Carlo methods. That is, TD learning can learn directly from interaction with the environment without any specific model of it (like Monte Carlo methods), and like DP, TD methods update it estimates without waiting for the final result (i.e., they perform incremental learning).

### 2.3.2.1 Dynamic Programming methods

Dynamic Programming offers a set of methods for solving the Bellman's equation (2.9) based on the Markov property. Two alternatives are found to affront this problem: policy iteration and value iteration.

[5] A reinforcement learning task that satisfies the Markov property is called a Markov decision process, or MDP. If the state and action spaces are finite, then it is called a finite Markov decision process (finite MDP).

Policy iteration

By this methodology, a policy is first evaluated and then improved. Evaluate a policy means to compute the value $V^{\pi}(s)$ for every state $s$ under policy $\pi$. Developing (2.9) and making use of the Markov property it can be written as

$$V^{\pi}(s) = R_s^{\pi(s)} + \gamma \sum_{s'} P_{ss'}^{\pi(s)} V^{\pi}(s'), \tag{2.12}$$

where $\pi(s)$ stands for the action taken under state $s$ and following policy $\pi$ (note that here is assumed that a deterministic action $a = \pi(s)$ is taken when arriving to a state), $R_s^a$ denotes the immediate reward of taking action $a$ under state $s$, and $s'$ denotes the following state. Note that (2.12) is a linear equation for $V^{\pi}(s')$ that can be solved by means of numerical methods. Then, the expected value of each state for each policy could be obtained provided that rewards and transitions probabilities are known.

Similarly (2.12) can be particularized for each action $a$ using values $Q^{\pi}(s,a)$:

$$Q^{\pi}(s,a) = R_s^a + \gamma \sum_{s'} P_{ss'}^a V^{\pi}(s'). \tag{2.13}$$

Once the expected values are computed, then policy improvement is applied. The idea is to find a policy $\pi^*$ that returns the largest values for each state. Then, maximizing the values over all possible actions will reveal if there is an action $a$ satisfying

$$\pi^*(s) = \arg\max_a \left( Q^{\pi}(s,a) \right), \tag{2.14}$$

and then $Q^{\pi}\left(s, \pi^*(s)\right) \geq Q^{\pi}\left(s, \pi(s)\right)$. The following pseudocode sums up the policy iteration procedure:

```
1.  choose an arbitrary policy π'
2.  loop
3.      π:=π'
4.      compute the value function of policy p:
5.          solve the linear equations
6.              V^π(s) = R_s^{π(s)} + γ Σ_{s'} P_{ss'}^{π(s)} V^π(s')
7.          improve the policy at each state:
8.              π*(s) = arg max_a ( Q^π(s,a) )
9.  until π=π'
```

Value iteration

Value iteration is the main alternative to policy iteration. Here the $Q$-values ($Q(s,a)$) are updated simultaneously according to the formula

$$Q'(s,a) = R_s^a + \gamma \sum_{s'} P_{ss'}^a \max_b Q(s',b), \tag{2.15}$$

where (2.15) differs from (2.13) in the values considered for the next state.

#### 2.3.2.2 Monte Carlo methods

In policy iteration and value iteration, the value function was calculated using the known transition probabilities and expected rewards. If those probabilities are not known, the value function can be learned. Given a fixed policy $\pi$ several complete iterations are performed. These iterations may execute a number of states, and after termination of a particular iteration, all the rewards for the states visited are known. Then, for each state $s$ visited during a particular iteration, $V^{\pi}(s)$ can be approximated by taken the average over the known future rewards. Because these methods are trial-based, they are sometimes referred to as "Monte Carlo" methods.

#### 2.3.2.3 Temporal-Difference methods

Temporal difference methods adjust the estimated value of a state based on the immediate reward and the estimated value of the next state instead of waiting to reward until the end of an episode.

Actor-Critic

It consists of two components: a *critic*, and a reinforcement-learning component or *actor*. The actor can be an instance of any of the k-armed bandit algorithms [26], modified to deal with multiple states and non-stationary rewards. But instead of acting to maximize instantaneous reward, it will be acting to maximize the heuristic value $v$ that is computed by the critic. The critic uses the real external reward signal to learn to map states to their expected discounted values.

The critic learns the value of a policy using Sutton's TD(0) algorithm [25]:

$$V(s) = V(s) + \alpha\left(r + \gamma V(s') - V(s)\right), \tag{2.16}$$

where $r$ is the instantaneous reward. The key idea is that $r + V(s')$ is a sample of the value of $V(s)$, and it is more likely to be correct because it incorporates the real $r$. If the learning rate $\alpha$ is adjusted properly (it must be slowly decreased) and the policy is held fixed, TD(0) is guaranteed to converge to the optimal value function.

Q-Learning

For a policy $\pi$, define Q-values (or action-values) as:

$$Q^{\pi}(s,a) = R_s^a + \gamma \sum_{s'} P_{ss'}^{\pi(s)} V^{\pi}(s') \tag{2.17}$$

In other words, the Q-value is the expected discounted reward for executing action $a$ at state $s$ and following policy $\pi$ thereafter. The object in Q-learning [27] is to estimate the Q-values for an optimal policy. For convenience, define these as $Q^*(s,a)$, for all $s$ and $a$. It is straightforward to show that

$$V^*(s) = \max_a Q^*(s,a), \tag{2.18}$$

and that if $a^*$ is an action at which the maximum is attained, then an optimal policy can be formed as $\pi^*(s) = a^*$. Herein lays the utility of the Q-values: if an agent can learn them, it can easily decide what it is optimal to do. Although there may be more than one optimal policy or $a^*$, the Q* values are unique.

In Q-learning, the agent's experience consists of a sequence of distinct stages or episodes. In the n-th episode, the agent observes its current state $s_n$, selects and performs an action $a_n$, observes the subsequent state $s'_n$, receives an immediate reward $r_n$ and adjusts its $Q_{n-1}$ values using a learning factor $\alpha_n$ according to:

$$Q_n(s,a) = \begin{cases} (1-\alpha_n)Q_{n-1}(s,a) + \alpha_n\left(r_n + \gamma V_{n-1}(s'_n)\right) & \text{if } x = x_n \text{ and } a = a_n \\ Q_{n-1}(s,a) & \text{otherwise} \end{cases} \tag{2.19}$$

where

$$V_{n-1}(s'_n) = \max_b (Q_{n-1}(s'_n, b)) \tag{2.20}$$

is the best the agent thinks it can do from state $s'_n$. Of course, in the early stages of learning, the Q-values may not accurately reflect the policy they implicitly define. The initial Q-values, $Q_0(s, a)$, for all states and actions are assumed to be given [27].

## 2.3.3 REINFORCE methods

Reinforcement Learning algorithms proposed in Chapter 5 are based on the REINFORCE algorithms originally presented by Williams in [28], and modified later by Phansalkar and Thathachar in [29] to give REINFORCE the ability of performing global optimization of the reward signal thanks to the inclusion of a perturbation term. REINFORCE (REward Increment = Nonnegative Factor times Offset Reinforcement times Characteristic Eligibility) algorithms are a set of simple algorithms of reinforcement learning for associative learning (the learner is required to perform an input-output mapping) and connectionist networks. They involve immediate reward (the payoff is obtained from the last input-output mapping only). REINFORCE algorithms are based on the estimation of a relevant gradient and they do not store any information about the gradient estimation (the algorithms base on the climbing of an appropriate gradient). The REINFORCE algorithms are within the actor-critic methods of temporal-difference learning and were selected for the thesis work due to their inherent optimization features of a given reward signal and low implementation complexity. For instance, there is no need of storing Q-values as in Q-learning.

The input-output behavior has a random component that is used to perform the exploratory learning (search of new useful actions that improve the performance). In general there are multiple types of REINFORCE based units involving different types of probability distributions of the output. This thesis focuses on the simplest one based on Bernoulli distributions and logistic functions, receiving the name of *Bernoulli logistic units*.

### 2.3.3.1 Bernoulli logistic units

Consider a single RL agent $i$ represented in Figure 2.10, with a single input $x_i$ and a single output $y_i$. In general, the RL agent is described by the tuple $\langle R, X, Y, w_i, g_i, \Theta \rangle$. $R$ is the set of values $r_i$ that reward signal could take (in general any compact set of $\mathbb{R}$), which the agent tries to maximize in the long term. $X$ is a set of scalar context signals $x_i$ that reveal agent's environment actual status and can be inputs for the agent. In the same way, $Y$ is the set of actions of the agent, each one issued as an output value $y_i$. In general, the action chosen by the agent is a random variable, whose probability generating function is

denoted with $g_i$ and depends on both the internal status of the agent $w_i$ and the context signal $x_i$ provided. Finally, $\Theta$ is the learning algorithm presented in the next subsection.

RL agent's output for *Bernoulli-logistic units* (BLU) is a Bernoulli random variable $y_i \in Y = \{0,1\}$. That is, the probability generating function of the agent is of the form:

$$g_i(y_i, p_i) = \begin{cases} 1 - p_i, & \text{if } y_i = 0; \\ p_i, & \text{if } y_i = 1. \end{cases} \tag{2.21}$$

That is, it is a single parameter probability distribution, where the output of the agent has associated a certain action-selection probability $p_i$. One can consider then that $p_i$ contains *knowledge* of the agent since it represents how often each one of the two actions is selected. Agent's input $x_i$ and internal status $w_i$ are related with $p_i$ by means of the logistic function as

$$p_i = f(x_i w_i) = \frac{1}{1 + e^{-x_i w_i}}. \tag{2.22}$$

The logistic function is a differentiable function whose shape determines how fast the internal probability tends to 0 or 1, then determining the action that the agent will mostly select. Figure 2.11 illustrates the shape of the function respect the input value $z_i = x_i w_i$. It can be seen that for input values between -5 and 5 the logistic quickly tends to 0 or 1 respectively.

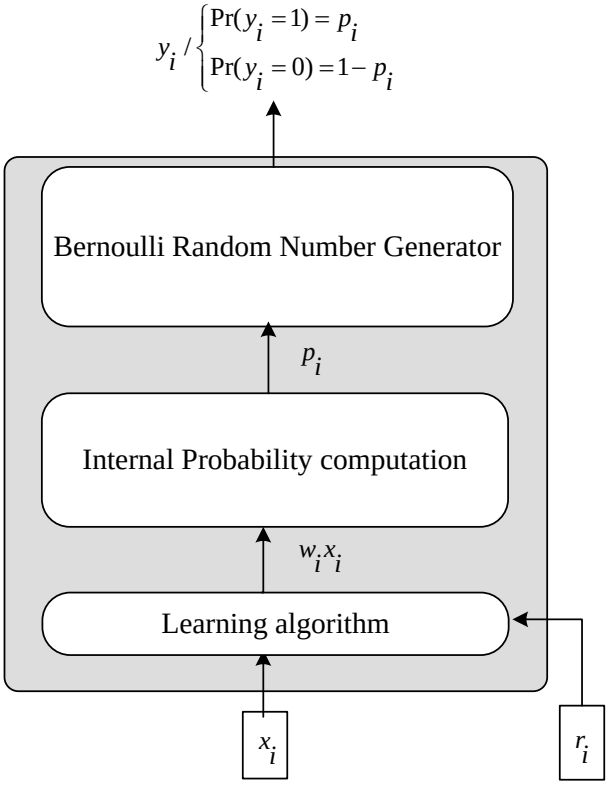


**Figure 2.10 General scheme of a REINFORCE RL agent**

As shown in the following, the learning algorithm will enforce the reward received in the internal probabilities, thus, affecting the frequency that each one of the two actions is selected.

### 2.3.3.2 Algorithm description

Consider that the agent interacts with an environment in a succession of time steps (denoted with $t$), where for each action $y_i(t)$ a reward signal $r_i(t)$ is returned to the agent. The learning capability of the agent is condensed step-by-step in the internal status $w_i(t)$, which is updated in accordance with the following learning rule [29]:

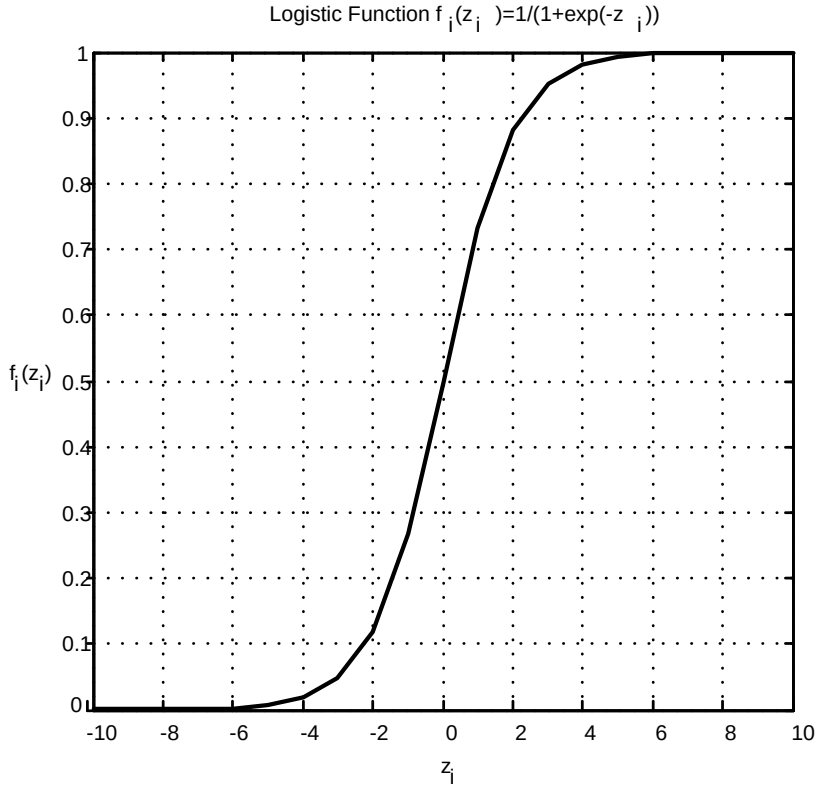


**Figure 2.11 Logistic function shape**

$$w_i(t) = w_i(t-1) + \Delta w_i(t)\,, \tag{2.23}$$

$$\Delta w_i(t) = \alpha(t)\left(r_i(t) - \overline{r}_i(t-1)\right)\frac{\partial \ln g_i\left(y_i(t-1), x_i(t-1), w_i(t-1)\right)}{\partial w_i} + \alpha(t)\xi\left(w_i(t-1)\right) + \sqrt{\alpha(t)}\zeta_i(t)\,. \tag{2.24}$$

Note that, by varying the internal status, $w_i(t)$, the learning algorithm is varying the selection probabilities in (2.22) so knowledge acquired from current reward is certainly enforced in the agent. Learning rule in (2.24) is compound of three terms.

The first term in (2.24) performs the gradient ascent of the reward signal in the particular case of Bernoulli logistic units considered here. Parameter $\alpha(t) > 0$ is called the learning rate, which is decreased with the RL steps to improve the convergence behavior of the algorithm [30]. More precisely, the learning rate is linearly decreased as $\alpha(t) = \alpha(t-1) - \Delta$, where $\Delta$ is a factor small enough to assure a smooth transition between steps and to guarantee that the learning rate is always positive within a given number of steps. $\overline{r}_i(t)$ is the reinforcement baseline or average reward computed from the reinforcement signal as

$$\overline{r}_i(t) = \beta r_i(t) + (1-\beta)\overline{r}_i(t-1), \tag{2.25}$$

with $0 < \beta \leq 1$. $\partial \ln g_i / \partial w_i$ is called the characteristic eligibility of $w_i$ and by using (2.21) and (2.22), can be written in terms of the current output, internal probability, and context input as

$$\frac{\partial \ln g_i\left(y_i(t), x_i(t), w_i(t)\right)}{\partial w_i} = \left(y_i(t) - p_i(t)\right) x_i(t), \tag{2.26}$$

which reduces the derivative computation to a simple subtraction and multiplication. This term would be enough to optimize the expected value of the reward signal denoted as $E\{r \mid \mathbf{W}\}$, where $\mathbf{W}$ is the weight matrix of the RL system composed by all the weighting values of several RL agents. Concretely, it is shown in [28] that the average update vector $E\{\Delta\mathbf{W} \mid \mathbf{W}\}$ is proportional to $\nabla_{\mathbf{W}} E\{r \mid \mathbf{W}\}$, the gradient of the average reinforcement signal. Then, following the first term in (2.24) to update the weights, the weight matrix found when the algorithm converges (i.e. $\Delta\mathbf{W} = 0$) maximizes $E\{r \mid \mathbf{W}\}$, that is, first term in (2.24) makes the weighting matrix lie in the weight space in a direction for which the performance measure is increasing. However, it was shown in [29] that this was not sufficient to assure *global* maximization of the reward signal. Thus, the second and third terms included in (2.24) were proposed.

The second term in (2.24) bounds the behavior of the algorithm and introduces a minimum exploratory behavior in the agent, with a small exploratory probability, $0 < p_{\text{explore}} \ll 1$, through the following function

$$\xi\left(w_i(t)\right) = \begin{cases} \Omega - w_i(t), & w_i(t) \geq \Omega; \\ 0, & \left|w_i(t)\right| < \Omega; \\ -\Omega - w_i(t), & w_i(t) \leq -\Omega; \end{cases} \tag{2.27}$$

where $\Omega = \ln\left((1 - p_{\text{explore}})/p_{\text{explore}}\right)$. This probability is necessary to allow the algorithm to explore new actions not taken in the past in order to seek for better reward response.

Finally, the third term introduces a perturbation parameter $\zeta_i(t)$, which is a random variable of zero mean and variance $\sigma^2$ (e.g., in this paper $\zeta_i(t)$ takes the value either $+\sigma$ or $-\sigma$ with equal probability, being $\sigma$ a positive constant). This term was proposed to give the algorithm the capability of escaping from local maxima and reaching global maximum of the average reward with a sufficient small value of $\sigma$ and a sufficient number of iterations of the learning loop [29].

## 2.4 Summary

This chapter has presented the technical concepts where this thesis develops its work. First a description of an OFDMA radio interface has been presented, studying physical and link layer implementations and the standardized versions of OFDMA radio access in 3GPP LTE and WiMAX. Concepts such as OFDM and the typical time-frequency grid of OFDMA radio interfaces have been explained, denoting the high spectrum flexibility that this type of radio interface brings to 4G cellular mobile networks. Then, link layer procedures like dynamic packet scheduling and dynamic link adaptation were explained as essential mechanisms to provide multiuser access and reliable and variable data rate communications. It has also been illustrated how an OFDMA radio interface can combat

multipath propagation and fast frequency selective fading, providing an efficient usage of the radio resources thanks to multiuser diversity. Hence, these link layer procedures are essential in the short-term radio resource management of an OFDMA interface, and thus concrete implementation will be given in the thesis proposal for spectrum assignment framework for OFDMA radio interfaces in Chapter 4.

This chapter has also reviewed, in a second part, the basic concepts related to self-organization. Self-organization arises as an appealing concept to manage 4G mobile cellular networks. It is defined as the ability of a system to autonomously take decisions to orient its behavior for best adapting to a changing environment. Then, operational and capital expenditures reductions can be obtained for a 4G network thanks to the implementation of self-x (i.e., self-configuration, self-optimization, self-healing, etc.) procedures by following a self-x cycle. This cycle entails system status observation, system parameters optimization and system parameters re-configuration. As reflected in Chapter 4, those tasks have been included in the proposed spectrum assignment framework.

Finally, the third section of this chapter presented the foundation theory of reinforcement learning. This methodology has been selected to implement part of the algorithms proposed in Chapter 5 of this thesis. The objective is to exploit the intrinsic feature of reinforcement learning of optimizing a given reward function by following a cyclic interaction with an environment. The concepts and most relevant learning methods were presented, ending with the description of the particular method used in this thesis, called REINFORCE (REward Increment = Nonnegative Factor times Offset Reinforcement times Characteristic Eligibility) algorithms.

# 3

# LITERATURE REVIEW FOR SPECTRUM ASSIGNMENT STRATEGIES

## Outline

# 3 Literature review for spectrum assignment strategies

This chapter is devoted to review literature contributions in the ambit of spectrum assignment strategies on next generation cellular networks. It is worth to remark that the final aim of this chapter is to examine the different state-of-the-art proposals that have been used as a starting point for the main contributions of the thesis, that is, strategies for spectrum assignment in OFDMA networks, or the usage of self-organization and reinforcement learning in that context. Moreover, part of the proposals given in this chapter will be used as reference cases for the performance comparison results of this thesis.

In Section 3.1 some relevant strategies for frequency assignment in OFDMA networks are examined. First, several dynamic resource assignment strategies are presented and classified according to their functional architecture (centralized, decentralized) and their execution time scale (short-term, medium-term). Then, frequency planning strategies are reviewed, providing also a categorization of them. Finally, spectrum arrangement strategies in the context of two-layer deployments are presented, mainly focusing in cellular structures with macro- and femtocells.

In Section 3.2, relevant proposed spectrum assignment techniques for cellular networks based on self-organization and reinforcement learning are examined.

## 3.1 Proposals for Frequency assignment in OFDMA networks

Chapter 2 highlighted the spectral flexibility of an OFDMA radio interface where the high granularity in both time and frequency (i.e., the time-frequency grid) facilitates the definition of radio resource management strategies, like packet scheduling, that dynamically assign radio resources to users. However, this spectral flexibility can be extended to the cell frequency assignment problem in a multicell OFDMA mobile cellular network as described in this clause.

Being the interference management a key issue in all the proposals for frequency assignment in OFDMA networks, in the sequel a simplified mobile cellular network showing the different interference sources is introduced. Figure 3.1 illustrates a mobile cellular network with a downlink radio interface based on OFDMA. The mobile network operator covers the whole service area with several base stations that cover smaller areas (cells). In general, the interference in cellular systems can be classified between *intracell interference* when it takes place between transmissions from/to users served by the same cell and *intercell interference* when it affects to transmission from/to users served by different cells. An OFDMA radio interface is usually free of intracell interference because packet scheduling mechanisms only allow that a resource block (i.e., a given chunk in a given frame of the time-frequency grid) can be assigned to a single user within a given cell at a given time. On the other hand, due to spatial and temporal traffic variations and the lack of coordination between the traffic schedulers in different cells, intercell interference cannot be easily avoided with scheduling mechanisms.

Intercell interference can degrade the performance of an OFDMA cellular system because:

- It provokes a reduction of the signal to interference plus noise ratio per resource block, what supposes that modulation and coding schemes of lower spectral efficiency and higher error protection have to be selected, hence lowering the amount of useful bits that can be sent per resource block. In other words, the user throughput in bits/s decays.

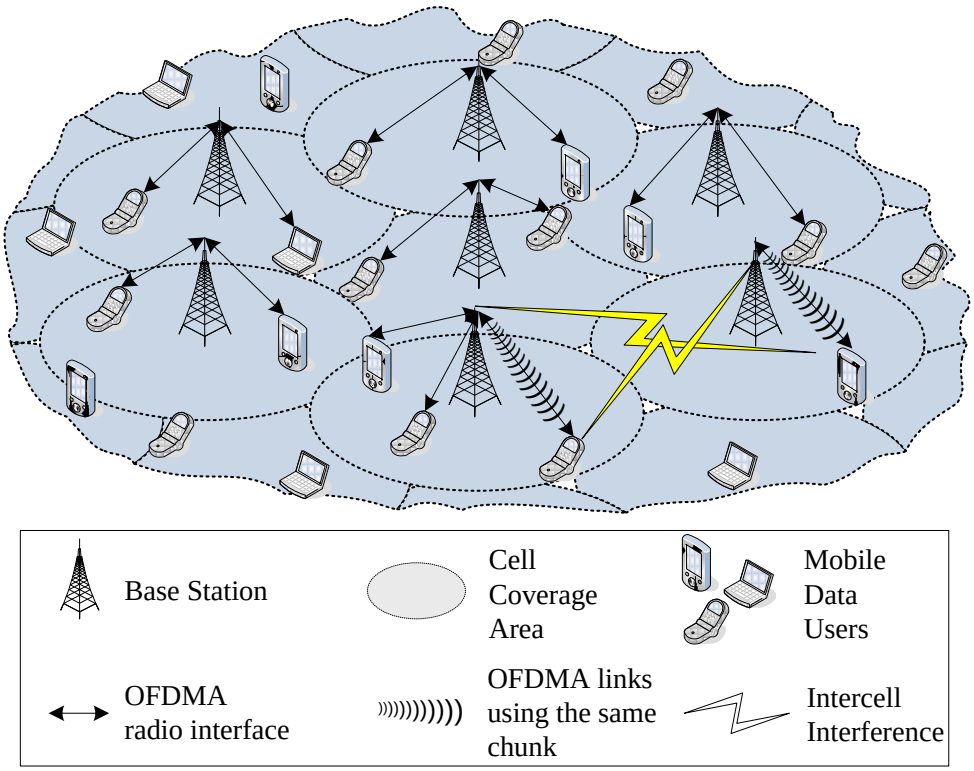


**Figure 3.1 Illustration of a mobile cellular network**

- In general, more retransmissions are needed to successfully deliver data packets, therefore increasing the signaling load.
- Transmission power consumption increases to assure a correct reception of the signal. This is particularly relevant for the uplink.

Therefore, in the following several strategies to reduce intercell interference in OFDMA-based cellular networks are examined.

### 3.1.1 Dynamic resource assignment

Radio Resource Management (RRM) algorithms for OFDMA based cellular networks [12] cover different aspects mainly focused in the resource assignment to users and cells. The key idea is how to assign subchannels[6], power, and modulation schemes to users and cells in an optimal or sub-optimal way. Among the optimization objectives, there are several possibilities. Chieochan and Hossain [31] presented an overview of current trends, dividing the strategies between those that minimize power consumption, those that maximize the assigned throughput, and those that optimize a given utility function, such as fairness in quality of service (QoS) fulfillment between users.

Approaches that minimize power consumption address the problem of assigning subchannels and bits to point-to-point wireless links in the presence of cochannel interference and Rayleigh fading. The objective is to minimize the total transmitted power over the entire network while satisfying the data rate

[6] In the following the word "subchannel" is used to refer to either a single subcarrier or a group of contiguous subcarriers (i.e., chunks defined in Chapter 2).

requirement of each link. Pietrzyk and Janssen presented in [32] a centralized heuristic optimization algorithm in a multicell scenario where OFDMA subchannels are reused between cells. Cochannel interference is controlled and QoS maintained for existing connections. With these characteristics, the algorithm presented is a scheduling algorithm where resources are assigned to each user and cell in short-term. Consequently, accurate knowledge of the channel state information is needed. On the other hand, Kulkarni, Adlakha, and Srivastava [33] proposed a decentralized algorithm to avoid that a centralized entity has to manage and collect information of the whole network.

Other approaches aim at maximizing overall system throughput [34][35], usually taking into account the intercell interference and certain power constraints. For instance, [34] makes use of *waterfilling* techniques to optimally distribute power between subchannels at the same time that data rate per subchannel is maximized.

Finally, there are proposals that focus on the assurance of certain fairness guarantees between the throughput attained by the users in the central and edge areas of the cell [36][37]. Heo, Cha, and Chang [36] introduce a subchannel assignment algorithm in the short-term based on SINR measurements in downlink. Fairness is pursued by selecting the users that present the highest quotient between the achievable rate and the granted rate in the past. Additionally, the proposed strategy also minimizes the power devoted to each user. Remaining power is distributed among the users with the worst channel condition. The mechanism does not maximize throughput but improve fairness. Jorguseski and Prasad [37] present key aspects of different subchannel assignment strategies on downlink OFDMA systems. They present performance results in terms of throughput, and it is concluded that strategies that fully take into account current channel state information would get the best performance in average but would incur in unfair assignments for users in the cell's edge.

For all the approaches above, great computational efforts and information exchange among users and cells is needed to perform the subchannel-power-modulation-cell-user assignment in the short-term, mainly due to instantaneous variations of subchannel conditions in multipath propagation environments. Therefore, hierarchical approaches [38][39] appear to reduce the signaling overhead. These hierarchical schemes facilitate the medium-term cell-by-cell spectrum assignments that adapt to temporal and spatial variations of the load, and open the possibility to create cognitive secondary spectrum usage opportunities. On the other hand, the short-term assignment of radio resources to users is independently provided at cell level.

Li and Liu [38] proposed to perform the resource assignment in two temporal scales: in the medium-term the controller of a cluster of cells decides which subchannels should be used by each cell under control, trying to maximize the total cell's throughput and to perform intercell interference avoidance. In the short-term, the cell decides how to schedule users' transmissions into available subchannels regarding Channel State Information (CSI) reported by the users. It is a semi-distributed approach that shows low signaling overhead but only optimizes throughput without taking into account user's QoS requirements. Kwon et al. [39] also propose an assignment scheme divided in two stages that guarantees minimum QoS for users' sessions. Then, different subchannels are assigned to different cells depending on users' QoS requirements per cell. However, the problem is solved for a virtual cell composed of adjacent sectors of adjacent cells and does not take into account interference from farther co-channel sectors.

### 3.1.2 Frequency planning strategies

Several schemes regarding frequency planning[7] and intercell mitigation have been envisaged for multicell OFDMA networks. Due to excessive intercell interference, especially from adjacent cells, some users at the edge could not be served. Thus, Frequency Reuse Factor (FRF) concept was introduced leading to frequency planning schemes where intercell interference is highly reduced, because orthogonal subsets of subchannels are distributed among cells [40].

The simplest FRF scheme is FRF=1 where all chunks are available at any cell and are transmitted with the same power (denoted in the following as FRF1). In this scheme the users at the edge of the cell experience high interference because all neighboring cells reuse the same frequency. A possibility to mitigate this interference for these users is to increase the reuse factor. Applying a FRF $= M > 1$ implies that the total band is divided into $M$ equal subbands, and is distributed over groups of $M$ contiguous cells (or clusters) repeating this pattern in the cellular system. As in the FRF1, all chunks are transmitted with the same power. This scheme reduces the intercell interference but reduces by $M$ the cell potential capacity as well. Typically, FRF3 and FRF7 are classical frequency planning schemes, where the entire bandwidth is distributed among clusters of $M$ =3 and $M$ =7 cells respectively. Figure 3.2 illustrates the FRF for $M$ =3.

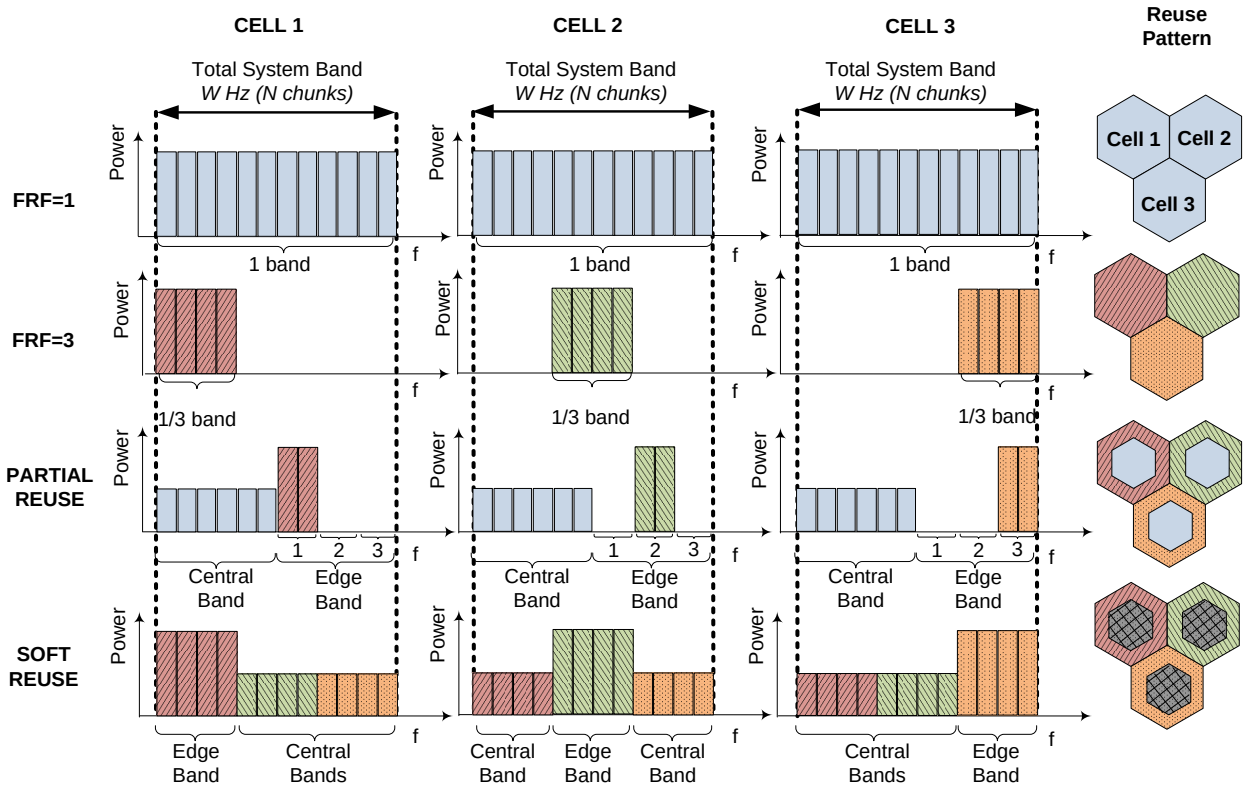


**Figure 3.2 Frequency planning schemes for OFDMA radio interfaces**

[7] Frequency planning tasks involve the assignment of spectrum resources to *cells* off-line during network deployment usually in a regular basis and taking into account the intercell interference. Then, the spectrum assignment remains unaltered until new infrastructure (i.e., cell sites) is added to the system, when a tedious manual and human supervised frequency planning is repeated.

Nevertheless, FRF schemes considerably reduce cell capacity. Mustonen [41] showed that global network performance can be improved by fragmenting spectrum according to the distance of the users to base stations. Hence, there are also other reuse schemes that are based on the division of the frequency band, the users, and the cells between central (inner) and edge (outer) sets. Figure 3.3 depicts an example of the cellular deployment for these schemes. Users within a cell are classified between central or edge users depending on whether they are located in the inner or outer part of the cell respectively. $C$ chunks are reserved for the central subband and $E$ chunks for the edge subband. The inner zone radius $D$ is set so that the percentage of the total cell area devoted to the inner part equals the percentage of cell chunks $\rho$ for that part.

Central chunks are transmitted with power $P_C$ and edge chunks with power $P_E$ ($P_C \leq P_E$). Concretely, the power per edge chunk $P_E$ is $P_E = P_{\max}$ to assure coverage in the whole cell. On the other hand, the power of each central chunk $P_C$ is adjusted in order to assure that average Signal to Interference Ratio (SIR) experienced by a central user located at a distance $D$ from the center of the cell is above a protection target $\gamma_0$ whenever the worst interfering cell is using the same chunk for edge usage (see $SIR_{E\to C}$ in Figure 3.3).

Therefore, due to coverage reasons, central chunks are for exclusive usage of central users whereas edge users have priority to use edge chunks. However, central users may be granted with the edge chunks if edge users' transmissions are not scheduled in them. Notice that the worst Signal to Interference Ratio (denoted as $SIR_{E\to C}$) for central users in a cell is given when the assigned chunks for central usage in this cell are the same as those devoted to edge usage in neighboring cells. Conversely, in the same conditions edge users increment their SIR ($SIR_{C\to E}$) because neighboring cells transmit with power $P_C \leq P_E$.

There several proposals to arrange the division between the central and edge subbands, where the two most representative, called *Partial-frequency Reuse* and *Soft-frequency Reuse* respectively are illustrated in Figure 3.2 and explained in the following. On the one hand, Partial-frequency Reuse (PR) schemes that

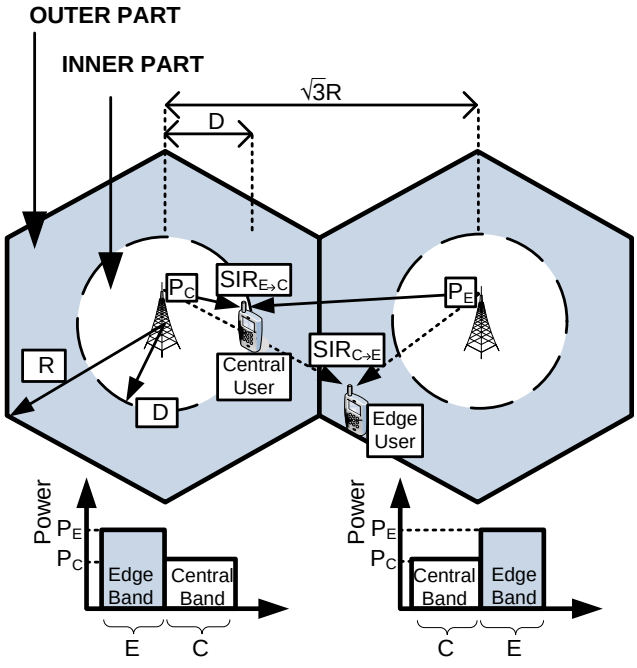


**Figure 3.3 Hybrid frequency reuse factor deployment**

divide the entire frequency band of the system between a central and an edge subband have been proposed [42][43]. Usually, the central subband is available in all cells (FRF1) whereas the edge subband is further divided in three equal subsets that are distributed regularly (FRF3) over cells. Hence, in this case $C + E/3$ chunks are assigned per cell, that is, $\rho = C/(C + E/3)$ .

On the other hand, Soft-frequency Reuse (SR) scheme [44] divides the frequency band into three subbands, all of them available in all cells (Figure 3.2). However, the edge subband serving the edge users, is transmitted with greater power than the other two central subbands that are only available for the central users. As in the FRF=1 scheme, all cells have all chunks available, but those belonging to the edge subband alternate their position into the system's band following a FRF3 scheme. The total number of chunks in each cell is $C + E$ chunks, that is, $\rho = C/(C + E)$ .

Moreover, dynamic management of the FRF has been proposed [45]-[48] where the FRF for each subchannel is dynamically adjusted to mitigate intercell interference and, in the end, increase system capacity guaranteeing users' QoS. Kim, Han and Koo [45] proposed optimal and suboptimal algorithms to dynamically adapt the frequency reuse factor of each OFDMA subchannel and maximize total system throughput while guaranteeing some user throughput QoS constrains. Then Choi, Kim and Bahk [46] introduced the Flexible FRF as an alternative to FRF1 and FRF3. However, the intermediate FRFs distribute the channels equally among cells, thus not being adapted to spatial and temporal variations of the network load. Lengoumbi, Godlewski and Martins [47] presented a rate optimization algorithm with minimum rates guarantees and interference mitigation. The problem is treated from a subchannel-cell-user assignment perspective. Finally, Elayoubi, Haddada, and Fourestie [48] proposed an analytical approach for both OFDMA downlink and uplink. Several FRF planning schemes are compared: FRF1, FRF3, PR, dynamic PR, where the power devoted to central and edge chunks is dynamically varied to adjust central and edge coverage.

Anyhow, all these FRF schemes impose several restrictions in the way that chunks are shared among cells since the number of chunks devoted to each FRF in each cell must be equal [49], which is not optimal for heterogeneous spatial distributions of the users (i.e., different traffic loads per cell) [50]. Furthermore, future trend in macrocell scenarios is that an intercell signaling interface between adjacent cells exists, so that it is possible to share information between them to perform distributed radio resource management tasks. Concretely, 3GPP LTE envisages the so called X2 interface, which enables the Inter-Cell Interference Coordination (ICIC) procedures [14].

Basically, in downlink, ICIC allows the coordinated scheduling of users' transmissions into chunks creating different FRFs between cells as those presented above. Figure 3.4 shows the functional scheme for the execution of ICIC in 3GPP LTE. The aim of ICIC strategies is to decide a set of restrictions and preferences over the usage and transmission power of the chunks in a given eNB, so that intercell interference is mitigated. These restrictions and preferences will be considered by the packet scheduling function when scheduling users' transmission in each eNB. Thus, by means of an adequate selection of the restrictions and preferences, it is possible to achieve certain coordination between the packet schedulers in different eNBs. To this end, certain interference indicators have been standardized. Regarding the downlink transmission, the *Relative Narrowband Transmit Power* (RNTP) indicator is shared through the X2 interface, which, roughly, anticipates the spectrum assignment that each cell is going to use in future. Hence, it can be considered that ICIC is a framework for developing decentralized spectrum assignment strategies, since each cell (eNB) independently executes this functionality.

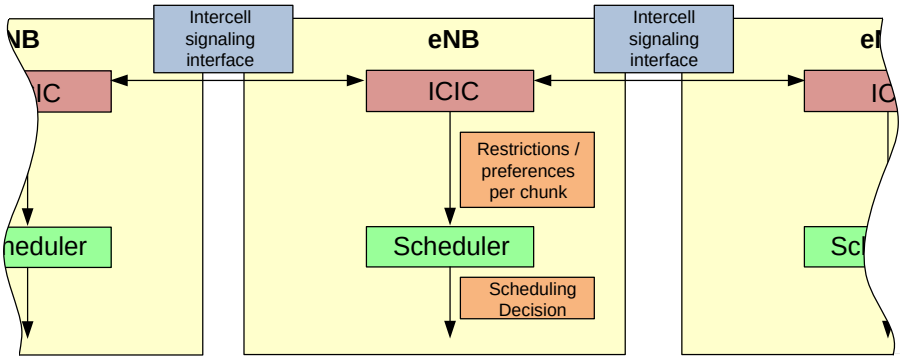


**Figure 3.4 ICIC mechanism functional scheme**

### 3.1.3 Spectrum arrangement in two-layer cellular deployments

Figure 3.5 depicts a two-layer cellular deployment. Typically, there is a layer of macrocell base stations with a regular hexagonal layout, where an intercell signaling interface could be available. Additionally, there is a layer of femtocell base stations that co-exists with the macrocell layer. Femtocells are small range user-deployed base stations introduced at a considerable amount of random locations [3]. They increase capacity, extend indoor coverage, and provide end-to-end communication through a DSL (Digital Subscriber Line) link.

The random and distributed nature of this new femtocell-based scenario would make difficult the success of a manual and centralized configuration of the spectrum assignment as usual in the macrocell deployment. Moreover, both macrocells and femtocells would coexist in most of the scenarios, which makes the spectrum assignment to reduce intercell interference even more challenging (i.e., intercell interference could appear between macrocells, between femtocells, and between macro- and femtocells). In the following, the different choices to arrange the spectrum in the two-layer cellular deployments are reviewed, especially focusing on decentralized approaches, where each cell (macro- or femtocell) determines its spectrum assignment.

Spectrum management in OFDMA femtocells is a challenging task where different possibilities are open [51]. In the presence of a layer of macrocell base stations over a layer of femtocells, one option is to divide the available spectrum into two separated subbands, one for each layer. This solution deploys an *orthogonal* spectrum assignment that completely eliminates the interference between the macrocell and femtocell layer, but limits also the capacity of each layer.

In contrast, *co-channel* spectrum assignment shares the available spectrum band between the macrocell and femtocell layers, increasing the available capacity for each one but at the cost of complex management of the interference. For instance, Espino and Markendahl [52] studied the interference generated by the femtocell layer in the macrocell deployment, determining the size of areas around femtocells where macrocell coverage decays dramatically. Moreover, Góra and Kolding [53] studied the cross layer interference patterns between macrocell and femtocells, for different femtocell deployments (sub-urban and dense urban) and different users' spatial distributions and traffic patterns, revealing a certain dependence with them. Additionally, Bai et al. [54] proposed a hybrid scheme where the spectrum access of the femtocell is performed to a co-channel or orthogonal part of the spectrum depending on the interference generated by the femtocell. Also Li et al. [55] proposed a flexible spectrum reuse scheme for femtocell deployment where a different channel reuse factor is employed according to where a femtocell

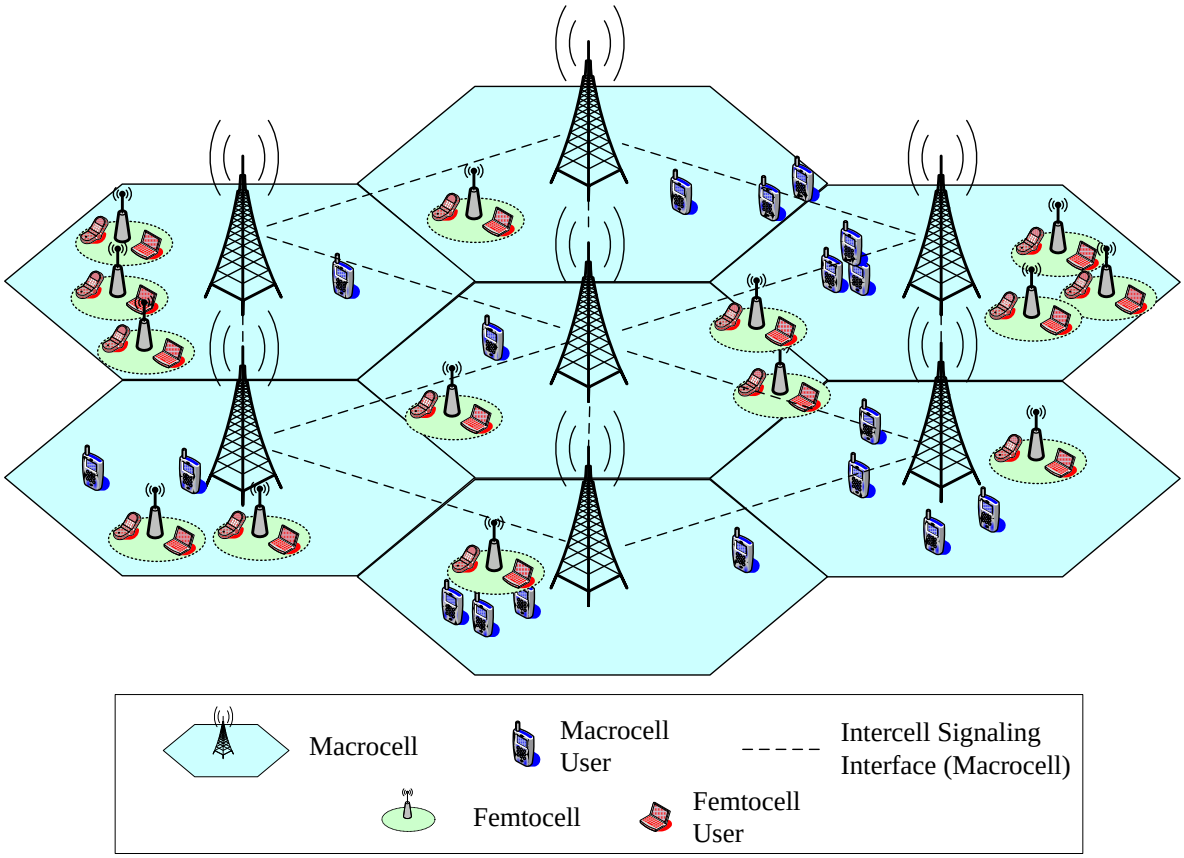


**Figure 3.5 Two-layer mobile cellular deployment**

is located in the macrocell. Finally, Claussen and Calin [56] showed that important capacity gains can be obtained in the macrocell layer by deriving indoor users to femtocell layer, where this gain depends on depends on the macrocell transmit power configuration affecting the interference produced to the femtocell layer.

In addition, femtocell spectrum management with regard to other femtocells is also an important task because in scenarios with a high density of femtocells, interference between them cannot be neglected due to their proximity. In this case, direct interactions between femtocells to coordinate their spectrum assignments are not, a priori, feasible, since an intercell signaling interface between femtocells could be quite challenging and, at the moment, is not envisaged in standardization bodies such us 3GPP LTE [14]. Hence, the reference approach in this case is a random access to a portion of the available spectrum [51][57]. For instance, the available band can be equally divided into $v$ portions, and each femtocell randomly selects one of them. Then a *Random*-FRF $v$ is implemented. The greater $v$, the lower the probability that two adjacent femtocells use the same portion, but also the lower the available capacity in each femtocell.

## 3.2 Proposals based on Self-organization and Reinforcement Learning

As described in Chapter 2 self-organization has recently paid a lot of attention in the context of mobile cellular networks. Usually, network observation and analysis of the network status are required to detect the instants when current spectrum assignment is no longer valid and then automatically trigger the mechanisms to manage the spectrum. Then, a new adequate spectrum assignment to the different

transmitters is decided. For instance, [51][58] propose self-organization schemes for frequency planning and coverage (power) in the context of femtocells networks, where self-organization takes a relevant role due to the distributed and autonomic nature of this scenario. Those references propose a *self-optimization loop* including *observation* of femtocell status, *analysis* and *decision* tasks. However, none of the references above include *learning* in the self-optimization procedure. In fact, continuous learning will improve the decision procedure through a suitable evaluation of the outcome resulting from prior decision under similar conditions.

Reinforcement Learning (RL) has been proposed for several applications in the field of mobile communications such as, radio resource management [59][60] or QoS provisioning [61][62]. In addition, RL has been successfully studied in a number of papers applied to dynamic channel assignment (DCA) in multi-cell FDMA networks [63]-[67], where the system bandwidth is divided into frequency channels that are assigned to incoming voice calls. In those references, comparison results with fixed channel assignment (FCA) and heuristic DCA schemes show that the RL approaches adopted improve the network's performance in terms of blocking probability while significantly reducing algorithms' complexity. Within the RL methods used to solve the DCA problem are Q-Learning [63][64][67], actor-critic TD(0) [66] and SARSA [65] (see Chapter 2 section 2.3 for details)

Nie and Haykin [63][64] propose an alternative method to solving the DCA problem based on Q-Learning. Their approach is designed to learn an optimal policy by directly interacting with the environment with which it works, that can be a real cellular network or a network simulator. Then learning is accomplished progressively by appropriately utilizing the past experience which is obtained during real-time operation and is stored (the Q-values) in lookup tables (memory consuming) or neural networks (memory efficiency but need of network training).

Similarly, Singh and Bertsekas [66] use TD(0) RL method to find the best system configurations learned from the past. Each time there is a new event in the network (call arrival or termination) the system explore the learned solutions to assign the one with the highest estimated value. To reduce and compact the states complexity, the proposed model uses a *feature extractor* module that delivers to the RL module reduced state information of the network. The reward bases on the number of ongoing calls.

Finally, Lilith and Dogancay [65] review aforementioned works and propose a modified RL model based on SARSA, which is an on-policy method since uses the same policy to estimate the next state-action value and to choice the next action. The most interesting result of this work is that reduced state RL models (i.e., models that do not take all possible problem states but a reduced wisely selected subset of states) may perform even better that full state models since the reduction in the total number of admissible state-action pairs means that each distinct state-action pair will be visited with a greater frequency over a given time (and consequently, convergence to an optimal policy is improved).

## 3.3 Summary

This chapter has presented a review of literature related to spectrum assignment in OFDMA networks and the usage of self-organization and reinforcement learning for that purpose. First, a state of the art with regard to dynamic resource assignment in multicell OFDMA networks has been presented. Two types of strategies have been found: on the one hand there are strategies that jointly solve the subchannel/power/cell/user assignment problem in the short-term. On the other hand, there are strategies

that solve independently the subchannel assignment to cells from the subchannel assignment to users. Spectrum assignment framework that will be presented in Chapter 4 will make use this latter approach. A summary of reviewed references can be found in Table 3.1.

Moreover, the subchannel assignment to cells based on the use of Frequency Reuse Factors (FRF) was reviewed. Several approaches were found, where proposals go from fixed deployment of FRFs to hybrid FRFs with different FRFs in the inner and outer part of the cell. Also dynamic management of the FRF has been found so that different FRF are used depending on intercell interference. Reviewed references are collected in Table 3.2. However those approaches impose several restrictions in the way that subchannels can be shared among cells. For instance, the number of subchannels devoted to each FRF in each cell must be equal, which is not optimal for heterogeneous spatial distribution of the users. Then, dynamic spectrum assignment algorithms given in this thesis in Chapter 5 will try to overcome this limitation.

The first part of this chapter ended with a review of contributions for spectrum assignment in the context of two-layer deployments. Concretely, coexistence between macro- and femtocell networks has been studied. Orthogonal and co-channel spectrum deployments have been presented, where the former devote different spectrum bands for the macro- and femtocell layer, avoiding cross-layer interference but reducing the available resources per layer. On the other hand, co-channel deployment shares the spectrum for both layers, at the cost of having cross-layer interference. Several references studying this issue have been presented. It is worth mentioning that approaches presented in that section are oriented for decentralized spectrum assignment strategies, where each cell decides its own spectrum assignment. Table 3.3 summarizes those references.

Finally, the second part of the chapter included an overview of literature proposals for spectrum assignment based on self-organization and reinforcement learning. Some references proposed a self-organization loop but learning was not considered. In fact, continuous learning will improve the decision procedure through a suitable evaluation of the outcome resulting from prior decision under similar conditions. Hence, Chapter 5 will propose RL algorithms to exploit this potential benefit. Table 3.4 summarizes the references for spectrum assignment considered.

**Table 3.1 Summary of dynamic resource assignment proposal in OFDMA networks**

| Ref | Authors | Title | Optimization Criterion | Architecture | Time scale |
|---|---|---|---|---|---|
| **[32]** | Pietrzyk, Janssen | Subcarrier allocation and power control for QoS provision in the presence of CCI for the downlink of cellular OFDMA systems | Power | Centralized | Short-term |
| **[33]** | Kulkarni, Adlakha, Srivastava | Subcarrier Allocation and Bit Loading Algorithms for OFDMA-Based Wireless Networks | Power | Centralized/ Decentralized | Short-term |
| **[34]** | Kim, Kim, Han | Subcarrier and Power Allocation in OFDMA Systems | Rate | Centralized | Short-term |
| **[35]** | Yan, Wenan, Junde | An Adaptive Subcarrier, Bit and Power Allocation Algorithm for multiuser ODFM systems | Rate | Centralized | Short-term |

| | | | | | |
|---|---|---|---|---|---|
| **[36]** | Heo, Cha, Chang | Effective adaptive transmit power allocation algorithm considering dynamic channel allocation in reuse partitioning-based OFDMA system | Fairness | Centralized | Short-term |
| **[37]** | Jorguseski, Prasad | Downlink Resource Allocation in Beyond 3G OFDMA cellular systems | Fairness | Centralized | Short-term |
| **[38]** | Li, Liu | Downlink Radio Rasource Allocation for Multi-Cell OFDMA System | Rate | Hierarchical | Short-term/Long term |
| **[39]** | Kwon, Lee, Gi Lee | Low-Overhead Resource Allocation with Load Balancing in Multi-cell OFDMA Systems | Rate | Hierarchical | Short-term/Long term |

**Table 3.2 Summary of OFDMA frequency planning proposals**

| Ref | Authors | Title | Frequency planning | Compared schemes | Comments |
|---|---|---|---|---|---|
| **[40]** | Wang, Stirling-Gallacher | Frequency reuse scheme for cellular OFDM systems | Fixed FRF | - | Concept definition |
| **[41]** | Mustonen, Hooli, Ylitalo, Tölli, | Application of Intra-RAN Flexible Spectrum Use to Networks with Multi-Rate Services | Variable sizes | - | Proofs that performance depends on the distance of users to cell |
| **[42]** | Sternad, Ottosson, Ahlen, Svensson | Attaining both Coverage and High Spectral Efficiency with Adaptive OFDM Downlinks | Hybrid FRF | PR | PR definition |
| **[43]** | Elayoubi, Fourestie | On frequency allocation in 3G LTE systems | Hybrid FRF | FRF1 FRF3 PR | - |
| **[44]** | Huawei | Soft Frequency Reuse Scheme for UTRAN LTE | Hybrid FRF | SR | - |
| **[45]** | Kim, Han, Koo | Optimal Subchannel Allocation Scheme in Multicell OFDMA Systems | Dynamic FRF | Optimal FRF1 FRF3 FRF7 | Proposes optimal OSAS and SAS algorithms |
| **[46]** | Choi, Kim, Bahk | Flexible Design of Frequency Reuse Factor in OFDMA Cellular Networks | Dynamic FRF | FRF1 FRF3 FRF7 | - |
| **[47]** | Lengoumbi Godlewski, Martins | Dynamic subcarrier Reuse with Rate Guaranty in a Downlink Multicell OFDMA system | Dynamic FRF | Several fixed FRF and heuristic algorithms | - |
| **[48]** | Elayoubi, Haddada, Fourestie, | Performance Evaluation of frequency Planning schemes in OFDMA-based Networks | Dynamic FRF | FRF1, FRF3, PR and power dynamic PR | Analytical studies |
| **[49]** | Xiang, Luo, Hartmann | Inter-cell Interference Mitigation through Flexible Resource Reuse in OFDMA based Communication Networks | Fixed and Hybrid FRF | FRF1 FRF3 PR SR | Performance comparison |
| **[50]** | López-Pérez, Jüttner, Zhang | Dynamic Frequency Planning Versus Frequency Re-use Schemes in OFDMA Networks | Fixed and Dynamic FRF | FRF1 FRF3 | - |

**Table 3.3 Summary of spectrum assignment proposals in two-layer deployments**

| Ref | Authors | Title | Target Layer | Scheme | Comments |
|---|---|---|---|---|---|
| **[51]** | Lopez-Perez, Valcarce, de la Roche; Zhang | OFDMA femtocells: a roadmap on interference avoidance | Macrocell Femtocell | Orthogonal Co-channel | Overview of spectrum assignment possibilities |
| **[52]** | Espino, Markendahl | Analysis of Macro - Femtocell Interference and Implications for Spectrum Allocation | Macrocell | Co-channel | Macrocell Coverage studies |
| **[53]** | Góra, Kolding | Deployment Aspects of 3G Femtocells | Macrocell | Co-channel | Cross-layer interference pattern in macrocell layer from femtocells |
| **[54]** | Bai, Zhou, Liu, Chen, Otsuka | Resource Coordination and Interference Mitigation between Macrocell and Femtocell | Macrocell Femtocell | Orthogonal Co-channel | Hybrid spectrum reuse |
| **[44]** | Li, Macuha, Sousa, Sato, Nanri | Cognitive Interference Management in 3G Femtocells | Femtocell | Co-channel | Different spectrum reuse depending on femtocell location |
| **[56]** | Claussen, Calin | Macrocell Offloading Benefits in Joint Macro and Femtocell Deployments | Macrocell | Co-channel | Macrocell capacity gains improvement |
| **[57]** | Chandrasekhar, Andrews | Spectrum allocation in tiered cellular networks | Macrocell Femtocell | Orthogonal Co-channel | Random Frequency reuse factor |

**Table 3.4 Summary of spectrum assignment proposals based on reinforcement learning**

| Ref | Authors | Title | Method | Comments |
|---|---|---|---|---|
| **[63]** | Nie, Haykin | A dynamic channel assignment policy through q-learning | Q-Learning | Solve DCA problem. Interaction with real cellular network or network simulator |
| **[64]** | Nie, Haykin | A q-learning-based dynamic channel assignment technique for mobile communication systems | Q-Learning | |
| **[65]** | Lilith Dogancay | Dynamic channel allocation for mobile cellular traffic using reduced-state reinforcement learning | SARSA | Reduced RL state model |
| **[66]** | Singh Bertsekas | Reinforcement learning for dynamic channel allocation in cellular telephone systems | TD(0) | Feature extractor model |
| **[67]** | Senouci Pujolle | Dynamic channel assignment in cellular networks: a reinforcement learning solution | Q-Learning | Multimedia traffic |

# 4

# FRAMEWORK FOR DYNAMIC SPECTRUM ASSIGNMENT

## Outline

# 4 Framework for Dynamic Spectrum Assignment

Chapter 3 described a number of strategies to perform radio resource assignment and interference mitigation in OFDMA-based radio access interfaces. From that study, it was concluded that certain strategies that perform the resource assignment simultaneously to users and cells could be prohibitively complex for practical implementation. Moreover, the usage of Frequency Reuse Factors to mitigate intercell interference limits the flexibility of the cellular network to adapt to heterogeneous spatial distributions of the traffic load. Hence, this chapter describes the framework proposed in this thesis to perform a dynamic spectrum assignment in an OFDMA cellular deployment. It is dynamic in the sense that spectrum assignment could change in time and space (cell-by-cell). The framework is based on self-organization to exploit the benefits of operational and expenditure reductions, adaptability, and constant performance improvement that, as pointed out in Chapters 2 and 3, this approach brings in the ambit of primary cellular networks and spectrum management.

## 4.1 Framework overview

The proposed framework is focused in a downlink radio interface based on OFDMA, where, as described in Chapter 2, radio resources are organized in a time-frequency grid divided into frames and frequency chunks respectively (see Figure 2.3 in Chapter 2). The aim of the framework is to provide the means to autonomously adapt the spectrum assignment per cell, in terms of assigned OFDMA chunks, depending on the spatial distribution of the traffic load (i.e., the traffic load per cell) and a given performance goal. This objective could vary, but usually will involve maximizing the spectral efficiency at the same time that a certain quality of service (QoS) is maintained. Also, other simultaneous targets could be considered such as mitigate intercell interference, or generate spectrum opportunities for secondary spectrum usage in a private commons spectrum access model [6].

The OFDMA resource assignment to users and cells is carried out in two temporal scales. This approach is adopted with the aim of reducing the resource assignment problem complexity of the short-term procedures [31] presented in Chapter 3. Hence, in the *short-term*, (i.e., at a frame scale of e.g. milliseconds) each cell is in charge of assigning radio resources to users by means of packet scheduling and link adaptation mechanisms (see Chapter 2 section 2.1.3). On the other hand, in the *medium-term*, (i.e., tens of seconds, minutes) a dynamic spectrum assignment controller (*DSA controller*) decides the next assignment of chunks to cells depending on the traffic load per cell and the performance objective. This DSA controller can be located in a single entity controlling the spectrum assignment for a set of cells, or being located at each cell, so that in this case, decisions of the DSA controller independently determine the spectrum assignment in each cell. Then, two approaches for the framework, *centralized* and *decentralized*, are proposed.

The centralized approach benefits from an overall picture of the network status and thus the centralized *Network DSA controller* can manage the spectrum assignment task by exactly knowing the impact that different spectrum assignments provoke between cells. Moreover, it is proven in this chapter that signaling to the network DSA controller is not a limiting factor. On the other hand, in the decentralized approach, a decentralized *Cell DSA controller* independently executes the spectrum assignment task at cell level giving certain robustness and scalability to the spectrum assignment process.

Nevertheless, the main benefit of the decentralized framework relies in its suitability for cellular deployments where a centralized entity is not envisaged, such as in femtocell networks composed by a large number of small access points. Figure 4.1 depicts an illustrative sketch for the centralized and decentralized approaches of the proposed DSA framework.

The key foundational stone for the network/cell DSA controller is *self-organization*, regardless the functional architecture framework (centralized or decentralized). Figure 4.2 illustrates the main self-organization concepts that are included in the proposed Network/Cell DSA controller [22]. *Observation* and *analysis* of the network/cell status are used to detect the instants when current spectrum assignment is no longer valid and then automatically *trigger* certain spectrum assignment strategies. Key Performance Indicators (KPI) are used to observe the status of the network/cell for a given assignment and to take the appropriate decisions. Some KPIs can be the average spectral efficiency, average QoS, etc.

Then, based on these KPIs, proper algorithms are executed in order to *decide* a new adequate spectrum assignment to the different transmitters. Moreover, continuous learning will improve the decision procedure through a suitable evaluation of the outcome resulting from prior decision under similar

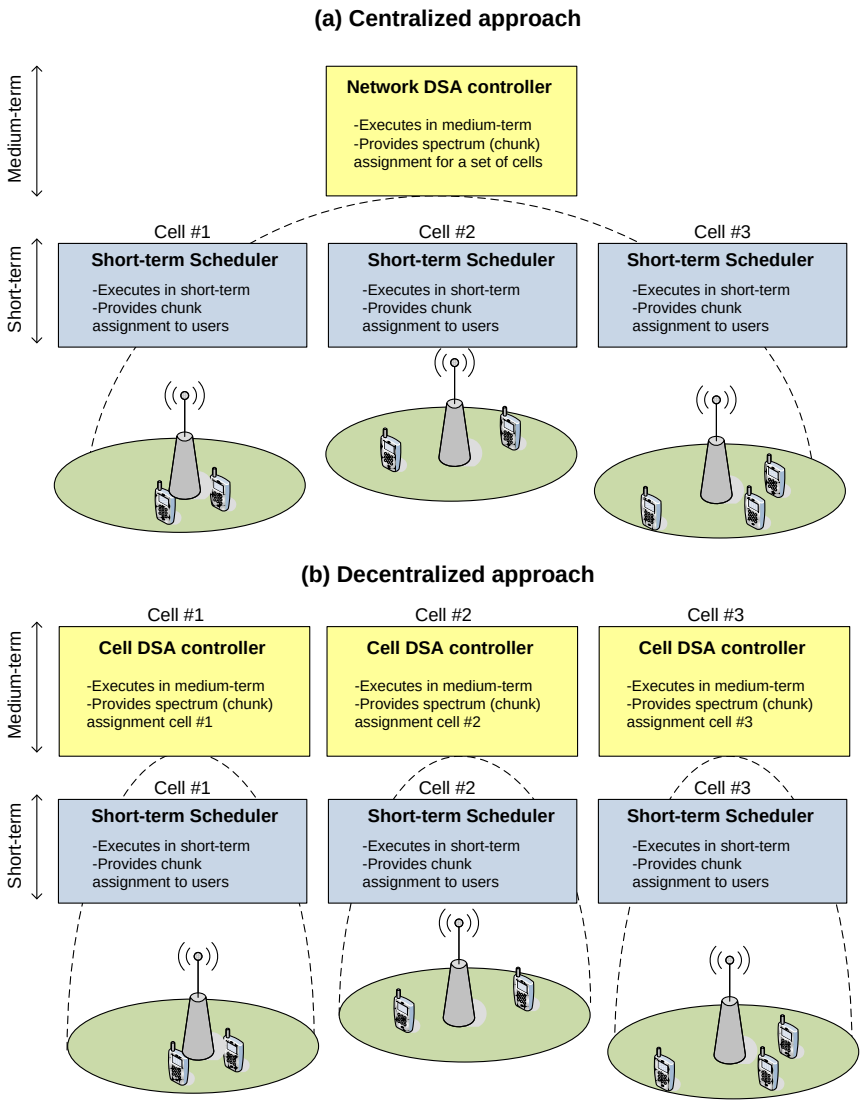


**Figure 4.1 Illustrative scheme of the centralized and decentralized approaches of the DSA framework**

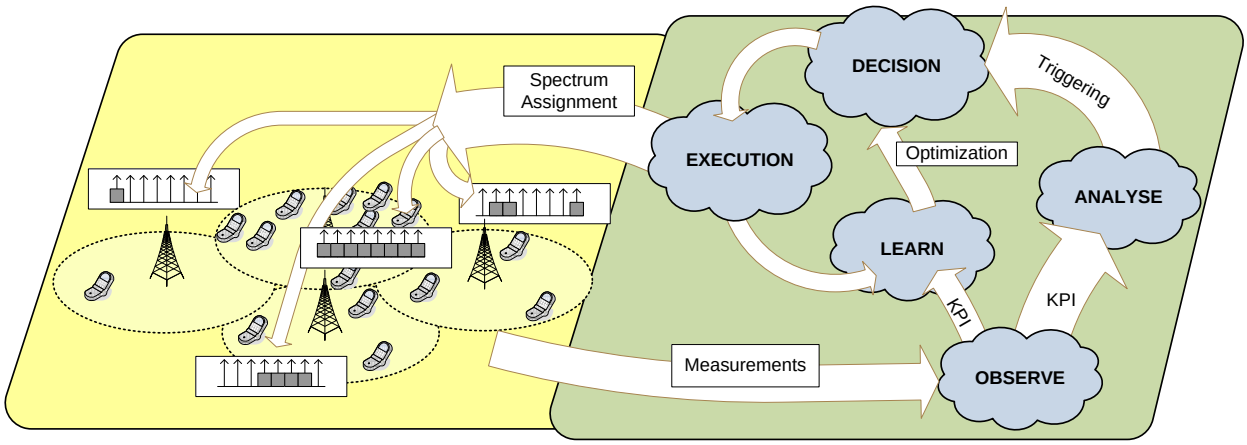


**Figure 4.2 Self-organization of spectrum in a cellular wireless network**

conditions. Finally, the execution phase provides the means to enforce the new spectrum assignment in the network/cell. It is worth mentioning that this self-organization approach makes this thesis proposal quite different from other approaches that usually face the spectrum assignment problem without exploiting the knowledge acquired from previous executions.

In the following clauses, first, the centralized version of the framework is presented, giving a detailed description of the functionalities associated to the short-term and medium-term execution. The role played by each one of the retained self-organization principles (observation and analysis, decision, learning and execution) is explained. Nevertheless, the specific details regarding the spectrum assignment algorithms will be given in Chapter 5. The interfaces and dynamics of the framework are also described here. Then, decentralized approach is explained in section 4.3, coinciding most of its functionalities with those implemented in the centralized approach. However, due to the decentralized functional architecture some procedures need to be redefined and special attention to the interaction between cells is paid.

## 4.2 Centralized framework

Figure 4.3 depicts a hierarchical architecture for an operator who deploys a multicell system with a downlink OFDMA-based radio interface. There is an uplink control channel to convey instantaneous (frame-by-frame) *measurements report* messages on the user side, from which the signal-to-interference plus noise ratio (SINR) in the different chunks can be obtained. These reports are necessary to perform the link adaptation for each established downlink communication in terms of an adequate modulation and coding rate scheme for a given SINR in a given frame.

As defined previously in Chapter 2 for an OFDMA radio interface, a total system bandwidth of $W$ Hz is divided into $N$ chunks. The minimum radio resource that can be assigned to a user in the short-term is a resource block, which is a chunk in a given frame. On the other hand, the spectral flexibility of the OFDMA interfaces allows for distributing chunks among cells, providing different spectrum assignments per cell to adapt to temporal and spatial distributions of the traffic load and improve spectral efficiency. The aim of the centralized framework is to provide the means to execute those tasks.

The *Short-Term Scheduler* (STS) in each cell is in charge of scheduling users' transmissions in the short-term (i.e., it assigns chunks to users) by using the available spectrum in the cell and performing

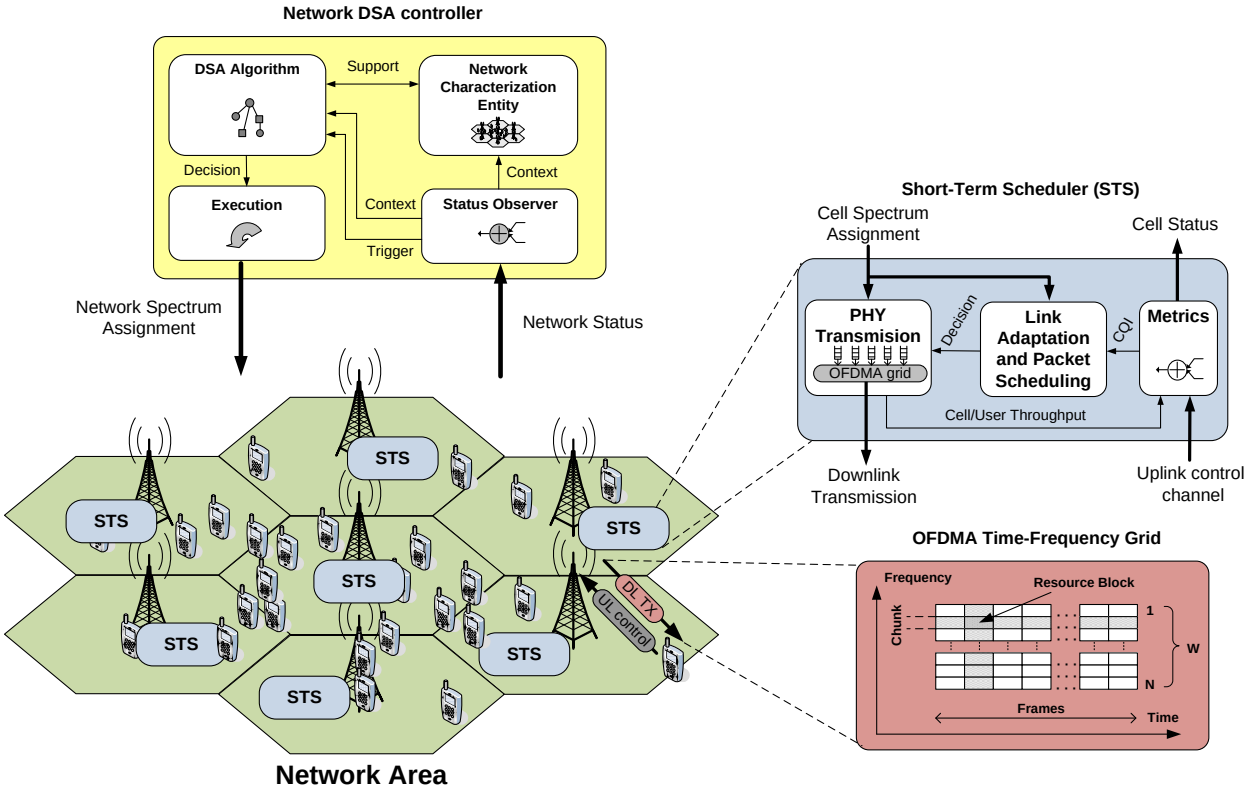


**Figure 4.3 Functional scheme of the centralized DSA framework**

OFDMA link adaptation. Also, the STS collects and averages different metrics to support the abovementioned procedures and establish cell status. STS details are given in section 4.2.1.

The *Network DSA controller* provides the cell-by-cell dynamic spectrum assignment in a set of cells in the medium-term by following self-organization principles. Roughly, the *Status Observer* decides when a new spectrum assignment is needed and provides the current network status (context). Then, the procedures to change the spectrum assignment for the set of cells under control of the Network DSA controller are started. It is the *DSA algorithm* who decides the new spectrum assignment. Two different methodologies have been employed in this thesis to build the DSA algorithm: heuristic algorithms and reinforcement learning algorithms. The *Network Characterization Entity* supports the execution of the DSA algorithm by providing an approximation to the response of the network for a given spectrum assignment selected by DSA algorithm. Finally, the Execution procedure provides the means to deploy the new spectrum assignment in the network.

Details regarding the Network DSA controller and its internal procedures are given in section 4.2.2 except for DSA algorithms which are extensively described in Chapter 5.

### 4.2.1 Short-term execution

Following the current trend of decentralizing functions towards edge network nodes, which enables shorter frame durations, lower latencies and higher speed channels, the Short-Term Scheduler (STS) is located at the base station of the cell, e.g. the eNB (E-UTRAN NodeB) as in the architecture proposed for Long-Term Evolution (LTE) by 3GPP in [14].

The objective of the STS is three-fold: (*i*) to collect and average necessary metrics to determine users' channel status and average cell status, (*ii*) to perform link adaptation, and (*iii*) to perform packet scheduling. All of them are explained in the following.

Metrics collection and averaging

The STS collects the necessary metrics to perform packet scheduling and link adaptation. This information can be retrieved from an uplink control channel where users convey their usual measurements reports messages. In a general view, the information contained in those messages can be an estimation of the channel quality in downlink, as the Signal to Interference plus Noise Ratio (SINR)[8]. The periodicity of the reported information is high enough so that packet scheduling and link adaptation can be performed in short-term (i.e., frame-by-frame). Moreover, the STS also supports Network DSA Controller execution by providing average *cell status*, whose content is detailed in section 4.2.2.

Link adaptation

STS also determines the user's modulation and coding rate in a frame-by-frame basis. Without loss of generality, it is considered that Channel Quality Indicator (CQI) for each user represents the SINR in each chunk, computed as:

$$\gamma_{m,n} = \frac{P_i G_{i,m} S_{i,m} F_{i,m,n}}{\sum_{\substack{j \in \Phi_n \\ j \neq i}} \left( P_j G_{j,m} S_{j,m} F_{j,m,n} \right) + P_{noise}}, \tag{4.1}$$

where $\gamma_{m,n}$ represents the SINR in the $n$-th chunk for the $m$-th user, index $i$ represents the serving cell and $j$ any interfering cell taken from the set of cells using the $n$-th chunk (denoted as $\Phi_n$). $P_i$ stands for the transmitted chunk power including transmitter and receiver antenna gains. Path loss and large scale fading is considered flat for all chunks while fast frequency selective fading may vary from one chunk to another depending on users' speed. $G_{i,m}$ denotes the distance dependant channel gain (inverse of path loss), $S_{i,m}$ the large scale fading (shadowing) and $F_{i,m,n}$ the fast frequency selective fading component that depends on the chunk $n$. Finally, $P_{noise}$ denotes the total thermal noise.

It is worth to remark that constant power per chunk has been considered since it has been concluded that non-constant power assignment brings marginal improvements [9]. Then, rate control is performed where users' instantaneous bit rate is varied according to the variations of the channel.

More precisely, the users' transmission bit rate is obtained as a function of the SINR through an Adaptive Modulation and Coding (AMC) scheme. There are different approaches in literature to perform this mapping. Here, analytic and table-based alternatives have been considered. Goldsmith and Chua propose in [68] the following formula for a target Bit Error Rate (BER) and Rayleigh fading channel:

$$q_{m,n}(f) = \log_2 \left( 1 + \frac{-1.5 \gamma_{m,n}(f)}{\ln(5BER)} \right), \tag{4.2}$$

[8] For instance, in the case of 3GPP LTE, the terminals are configured to send *channel status reports* (CSR) for all the bandwidth or a given subband. More precisely, the CSR contains the Channel Quality Indicator (CQI), which represents an index to a table indicating a modulation and coding scheme that the user suggests to use based on the cannel quality measured in downlink.

where $q_{m,n}(f)$ is the spectral efficiency in bits/s/Hz attained in the frame $f$, BER<(1/5)exp(-1.5) ≈ 4.46% and $\gamma_{m,n}(f)$ < 30 dB.

On the other hand, Schoenen, Halfmann, and Walke [69] propose the following table in the ambit of 3GPP LTE systems:

**Table 4.1 Adaptive modulation and coding**

| SINR threshold [dB] | Modulation *m* [bits/s/Hz] | Coding Rate *r* | Achievable spectral efficiency *q=m*r* [bits/s/Hz] |
|---|---|---|---|
| < 0.9 | - | - | 0 |
| ≥ 0.9 | 2 (QPSK) | 1/3 | 0.66 |
| ≥ 2.1 | 2 (QPSK) | 1/2 | 1 |
| ≥ 3.8 | 2 (QPSK) | 2/3 | 1.33 |
| ≥ 7.7 | 4 (16QAM) | 1/2 | 2 |
| ≥ 9.8 | 4 (16QAM) | 2/3 | 2.66 |
| ≥ 12.6 | 4 (16QAM) | 5/6 | 3.33 |
| ≥ 15.0 | 6 (64QAM) | 2/3 | 4 |
| ≥ 18.2 | 6 (64QAM) | 5/6 | 5 |

Packet scheduling

Finally, the STS performs packet scheduling to decide, frame-by-frame, which resource blocks of the OFDMA time-frequency grid are allocated to each user in given cell. Notice that due to DSA, a cell may not have all chunks available, and thus, the STS schedules transmissions only in the available resource blocks. The minimum radio resource block assignable to users is one chunk per frame, although more than one chunk can be assigned to a user in a given frame.

Different policies could be followed by the STS to perform packet scheduling. For example, they may be scheduled as a function of channel quality, buffer delay, throughput, buffer occupancy, service, etc. Two broadly used strategies have been considered in this thesis: Round Robin (RR) and Proportional Fair (PF) [70]. The objective of these strategies is to decide, for each chunk $n$ and in each frame $f$ the user that will transmit, denoted with $m_n^*(f)$.

The Round Robin strategy simply assigns cyclically the channel to users that want to transmit without considering their channel status (i.e., it is a *traffic-aware* strategy, but not a *channel-aware* strategy). Hence, RR is a fair strategy in the sense that each user has the channel assigned for the same amount of time. However, RR is not adapted to users' requirements, because users with the worst channel condition will require the channel more time that those with better channel condition in order to have similar QoS.

Proportional Fair (PF) is a traffic- and channel-aware strategy that provides a good trade-off between fairness and throughput at the same time that exploits multiuser diversity in both time and frequency. PF selects the user to transmit as:

$$m_n^*(f) = \underbrace{\arg\max}_{m} \left\{ \frac{R_{m,n}(f)}{W_{m,n}(f)} \right\}, \tag{4.3}$$

where $R_{m,n}(f)$ represents the instantaneous achievable rate that user $m$ can get at chunk $n$ in case the chunk is assigned to him/her. That is $R_{m,n}(f) = (W/N)q(f)$, being $W/N$ the bandwidth of a chunk and $q_{m,n}(f)$ the achievable spectral efficiency due to AMC. $W_{m,n}(f)$ is the window-averaged version of $R_{m,n}(f)$ as follows:

$$W_{m,n}(f) = (1 - \frac{1}{T_W})W_{m,n}(f\text{-}1) + \frac{1}{T_W}\bar{R}_{m,n}(f\text{-}1), \tag{4.4}$$

where $T_W$ is the window size in frames and $\bar{R}_{m,n}(f)$ stands for the final received rate at frame $f$ after scheduling. That is:

$$\bar{R}_{m,n}(f) = \begin{cases} R_{m,n}(f) & \text{if channel assigned} \\ 0 & \text{otherwise} \end{cases} \tag{4.5}$$

Proportional Fair strategy can exploit the multiuser diversity because it is a channel-aware scheduler. Figure 4.4 shows the spectral efficiency in bits/s/Hz obtained per cell in case that RR or PF scheduling strategies are used. This spectral efficiency is obtained under different SINR patterns in the cell and for a typical urban pedestrian A channel model [71]. The maximum achievable spectral efficiency is 5 bits/s/Hz, corresponding to a 64 QAM modulation with a coding rate of 5/6. Observe that the spectral efficiency achieved by RR does not depend on the number of users, whereas for PF, spectral efficiency increases with the number of users. Hence, PF is certainly exploiting multiuser diversity. Notice that, beyond 10 users, the multiuser diversity gain of PF over RR remains constant, since, with 10 users is probable that at least one user has the best channel conditions. Finally, both strategies demonstrate the same performance for only one active user.

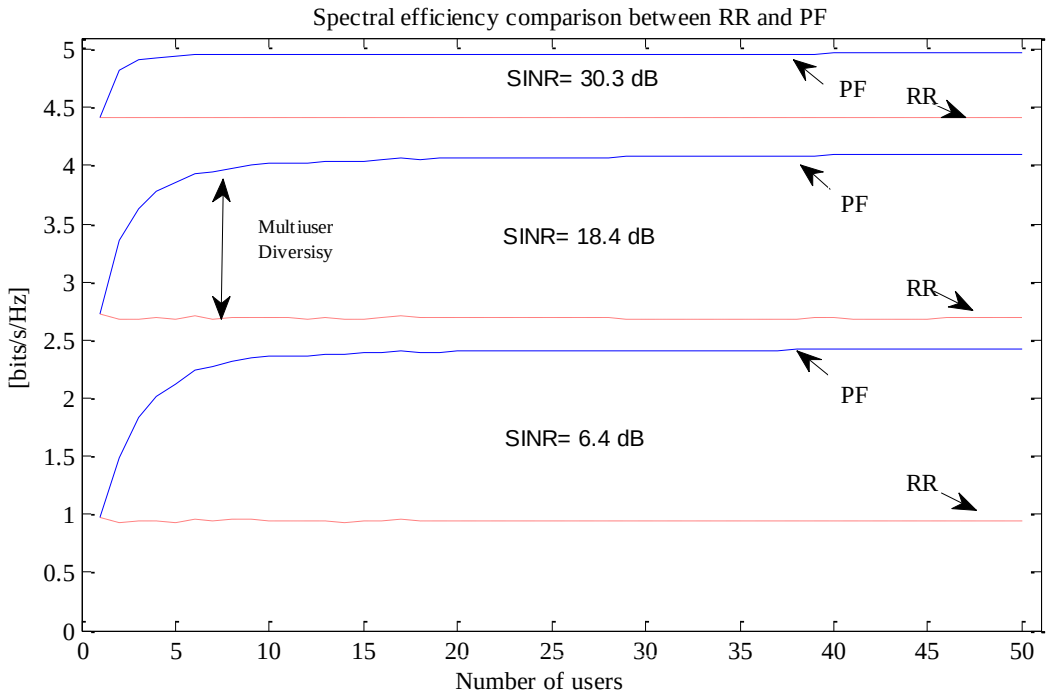


**Figure 4.4 Influence of the packet scheduling strategy in the exploitation of multiuser diversity**

## 4.2.2 Medium-term execution

Network DSA controller, which is located in a network node with the ability to control a set of cells, constitutes an intelligent entity with cognition capabilities to observe, analyze, decide, execute and learn. The objective of the Network DSA controller is to assign, in the medium-term, the right chunks per cell according to users' QoS needs.

Let consider $N$ available chunks numbered from 1 to $N$ in a downlink OFDMA cellular system, which have to be distributed over $K$ cells numbered from 1 to $K$. We define a $1\times KN$ assignment vector

$$\Upsilon = \left(u_{11},\cdots,u_{k(n-1)},u_{kn},u_{k(n+1)},\cdots,u_{KN}\right), \tag{4.6}$$

where $u_{kn}\in\{0,1\}$ denotes that $n$-th chunk is assigned to the $k$-th cell if $u_{kn}=1$ (not assigned if $u_{kn}=0$). The output of the Network DSA controller is a spectrum assignment vector that adapts the system's spectrum to traffic variations in time and space in the medium-term while maintaining users' satisfaction and avoiding intercell interference. Furthermore, it could attempt to achieve an efficient spectrum usage by releasing unnecessary frequency resources that can be used by secondary cognitive radio users in a private commons as seen in one of the application scenarios given in Chapter 6.

The functional blocks of the Network DSA controller are depicted in Figure 4.3 and described in the following.

### 4.2.2.1 Status Observer

The *Status Observer* entity *observes* and *analyzes* the network variable status. Then, it detects the instants when the current spectrum assignment is no longer valid to achieve a given users' QoS performance. Let $P_k^{th_{\text{target}}}$ be the average dissatisfaction probability per cell $k$ over a *measurements averaging period* of $l$ seconds, defined as the probability that the average user throughput is below a satisfaction throughput target $th_{\text{target}}$ (a formal definition of this metric will be given on Chapter 6). Based on this period, each cell reports to the Network DSA controller the dissatisfaction probability $P_k^{th_{\text{target}}}$ and the average number of users $U_k$. In fact, this constitutes the *cell status* build by the Short-Term Scheduler as mentioned in section 4.2.1.

The Status Observer computes average dissatisfaction probability for the network as

$$P^{th_{\text{target}}} = \frac{\sum_{k=1}^{K} U_k P_k^{th_{\text{target}}}}{\sum_{k=1}^{K} U_k}, \tag{4.7}$$

where $K$ is the number of cells of the network area. Then the Status Observer entity *triggers* the execution of the DSA algorithm if $P^{th_{\text{target}}}$ is either above a given threshold $\delta^{up}$ or below a threshold $\delta^{down}$, since in these cases current assigned resources would be either insufficient or over-provisioned, respectively, in accordance with the desired QoS.

In addition, Status Observer entity provides the *execution context* at the beginning of a new execution, which is the average number of active users per cell $U_k$ and the average dissatisfaction probability $P_k^{th_{\text{target}}}$.

As it will be seen in Chapter 5 this context orients the reasoning of DSA algorithm and allows the proposed algorithms to adapt to potential uneven distributions of the traffic load.

Finally, the required signaling exchange between the cells and the Network DSA controller is low since it is only produced on periods of $l$ seconds. Moreover, only few bytes are needed to encode the dissatisfaction probability $P_k^{th_{\text{target}}}$ and the average number of users $U_k$ per cell $k$, which act as inputs of the Status Observer entity. That is, if $P_k^{th_{\text{target}}}$ and $U_k$ are encoded with 8 bits (a maximum of 256 levels or users respectively), and one bit is devoted to determine if a specific chunk is assigned to a cell out of $N$ chunks (i.e., $KN$ bits encode the resultant spectrum assignment), then $K(N+8+8)$ bits are required to bear all the signaling for the execution of the DSA algorithm in a network area of $K$ cells. For instance, for a typical scenario composed of $K$=19 cells and $N$=12 chunks only 532 bits are needed, which is clearly bearable for current communication trunks of a network operator, taking into account that $l$ is in the order of seconds or tens of seconds.

#### 4.2.2.2 DSA algorithm

DSA algorithm implements the *decision* and *learning* functionalities and is in charge of optimizing the spectrum assignment in the set of cells depending on the required objectives (spectral efficiency, QoS, intercell interference mitigation, etc.). Its output is the next spectrum assignment for the network, in terms of the chunks that are assigned to each cell. It is worth to remark that the actual implementation of the algorithm is open in the framework. That is, one can consider different methodologies to implement the DSA algorithm in this functional block, as long as they provide as an output an assignment vector $\Upsilon$ as in (4.6).

Two methodologies have been used in this thesis to implement the DSA algorithm: heuristics and reinforcement learning. The heuristic approach is a simple but effective sub-optimal algorithm that estimates the number of chunks needed per cell depending on the traffic load in each cell. Then, the concrete chunks selected for each cell are decided based on intercell interference relations and the execution context. The reinforcement learning-based algorithms use reinforcement learning to decide a spectrum assignment per cell that maximizes a given reward metric appropriately defined in terms of spectral efficiency and QoS compliance. They are also simple to implement and exploit the intrinsic optimization features of reinforcement learning to seek the spectrum assignment vector $\Upsilon$. Implementation details for both approaches can be encountered in Chapter 5.

#### 4.2.2.3 Network Characterization Entity

Once the DSA algorithm is executed, it can request the support of a *Network Characterization Entity* (NCE). Basically, the NCE constitutes a *model* that approximates the network's response to a given spectrum assignment and execution context. For instance, for the heuristic algorithm, it provides the means to estimate the intercell interference under a given *candidate* spectrum assignment for solution. On the other hand, NCE could be useful for iterative algorithms such as reinforcement learning methods that run several candidate spectrum assignments in order to seek for the optimal one. Hence, NCE is a supporting entity that can be used by the DSA algorithm to learn how to orient its decisions.

Notice that NCE is an entity that the DSA algorithm consults off-line. This feature brings two major advantages over applying algorithms' actions directly to the live real network. First, the physical time taken by the real network to return a proper averaged response can be unacceptably slow in terms of

elapsed time compared with the quick response that NCE could provide. Second, candidate spectrum assignments can, in some cases, suppose a prohibitive cost in performance for the live network if they are applied on-line. On the contrary, applying those candidate spectrum assignments to an off-line entity only suppose a *simulation* cost, so, virtually, any candidate spectrum assignments can be tested.

How to build NCE depends on the performance objectives that one wants to capture in it (e.g., spectral efficiency, intercell-interference, etc.). Also, it is worth to highlight that, as for the DSA algorithm, NCE implementation is open and could depend on the needs of the DSA algorithm. For this reason, implementation details for the NCE procedures considered in this thesis are given at the end of Chapter 5 once the concrete implementation of the proposed DSA algorithms has been given.

#### 4.2.2.4 Execution Entity

The output of the DSA algorithm is a new spectrum assignment vector for the set of cells under control of the Network DSA controller. The Execution entity provides the means to actually deploy this spectrum assignment in the network. Then, it provides appropriate restrictions on the chunks available per cell. These restrictions have to be considered by the packet scheduling and link adaptation procedures in the Short-Term Scheduler, so that users' transmissions are only scheduled in the available chunks.

## 4.3 Decentralized framework

This section presents a decentralized approach for the spectrum assignment framework. Here, each cell behaves as an *autonomous* entity whose inputs are limited to *local measurements* and, optionally, *local interaction* with adjacent cells in case that an intercell signaling interface exists. This autonomous behavior of each cell makes the proposed approach suitable for any kind of cell and in particular for femtocells. Certainly, the distributed approach allows for much lower signaling load and flexibility than its centralized counterpart, which, on the other hand, would be unaffordable in e.g., femtocell scenarios.

Either distributed [72] or semi-distributed [73] approaches have been found in literature for spectrum assignment. In these cases, the dynamic spectrum assignment to cells and users in the short-term is performed, what leads to high computational requirements and requires a centralized coordinator to assist the spectrum assignment task across cells. However, in our proposed framework it is entirely distributed among cells and spectrum assignment is provided in medium-term, then considerably reducing execution time requirements and the signaling overhead.

The functional scheme of the decentralized framework is depicted in Figure 4.5 that shows an autonomous cell (macrocell or femtocell) surrounded of other cells. The framework focuses on a downlink OFDMA radio interface and there is an uplink control channel where users send measurements reports. Each cell performs autonomous spectrum assignment decisions with the objective of improving a certain objective (e.g., SINR) while guaranteeing cell users' QoS.

As in the centralized approach, the operation of the cell is divided into two timescales: in the short-term to handle users' traffic and perform OFDMA fast link adaptation, and in the medium-term (i.e., from tens of seconds to tens of minutes) to provide the spectrum assignment of the cell.

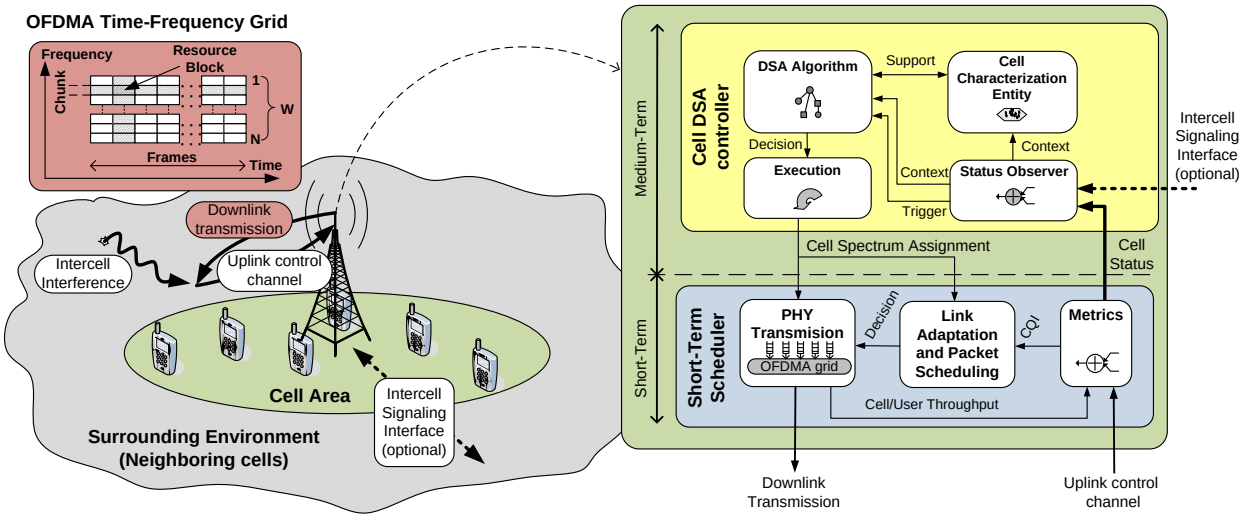


**Figure 4.5 Functional scheme of the decentralized DSA framework**

### 4.3.1 Short-term execution

The Short-Term Scheduler entity is devoted to collect users' measurements reports and average cell status, to perform short-term link adaptation and to execute packet scheduling mechanisms. Hence, Short-Term Scheduler functions are the same as in the centralized approach. Actually, the Short-term Scheduler entity was already envisaged as a "per cell" functional block in the centralized functional architecture. For this reason, implementation details can be found on section 4.2.1.

### 4.3.2 Medium-term execution

In the medium-term the cell determines which chunks it should use and which not by following a functional scheme, based on self-organization, very similar to the one explained for the centralized approach. Concretely, a *Cell DSA Controller* is included, whose aim is to find a suitable spectrum assignment that improves the average performance of the cell (e.g., SINR) while the QoS of the users in terms of a minimum assured throughput is assured. The Cell DSA controller is compound of the Status Observer, DSA Algorithm, Execution and the now called *Cell Characterization* entities, like the Network DSA controller in the centralized approach.

Nevertheless, due to the decentralized functional architecture, some particularities over the centralized approach arise such as the stability of the framework or the ways that cells employ to obtain information regarding the spectrum usage in nearby cells. In the following, the functional blocks of the Cell DSA controller are reviewed with emphasis in abovementioned particularities.

#### 4.3.2.1 Status Observer

The *Status Observer* entity in each cell is responsible of triggering the DSA algorithm in an *execution period* of $L$ seconds, which is common for all cells. This is different from the centralized approach where the execution was based on an event-based approach where a certain QoS metric (i.e., the dissatisfaction probability) was compared with given thresholds. The reason is for providing stability to the framework as explained in the following.

DSA algorithm is executed in a cell assuming that the spectrum assignment in adjacent cells is not varying during its execution period. Notice that, if the periods were aligned between cells (i.e., all cells execute DSA algorithm simultaneously), then the stability of the framework would be compromised, since in this case adjacent cells would be simultaneously varying their respective spectrum assignments. In order to avoid this issue, it is considered that, after switch-on, a cell randomly selects an initial time to execute the DSA algorithm for the first time, and then the execution period $L$ is followed from the initial time. Hence, since large values of $L$ are expected for a medium-term execution of the DSA algorithm, the probability that adjacent cells select the same initial time becomes negligible. For instance, taking $L$=60000 frames, leads to a probability of simultaneous execution of around $10^{-4}$ assuming a value of 6 adjacent cells[9].

In addition to that, the Status Observer entity collects and builds necessary inputs for the spectrum assignment decision task. Hence, local measurements to the cell are averaged over a *measurements averaging period* of $l$ seconds, where $l \ll L$ to favor that inputs at the time of execution reflect the latest up-to-date measurements. Then, the Status Observer provides to the DSA algorithm and the Cell Characterization Entity the average number of users in the cell, denoted as $U_k$, the average dissatisfaction probability per cell $P_k^{th_{\text{target}}}$, and the probability density function of the average SINR for each chunk ($\overline{\gamma}_k^{(n)}$), denoted as $s_{\overline{\gamma}_k^{(n)}}(\overline{\gamma}_k^{(n)})$. All, $U_k$, $P_k^{th_{\text{target}}}$ and $s_{\overline{\gamma}_k^{(n)}}(\overline{\gamma}_k^{(n)})$ constitute the cell status and execution context.

#### 4.3.2.2 DSA algorithm

The DSA algorithm implementation is in principle independent of the framework approach (centralized or decentralized). In fact, the output of the DSA algorithm should be an assignment vector

$$\Upsilon_k = \left(u_{k1}, \cdots, u_{k(n-1)}, u_{kn}, u_{k(n+1)}, \cdots, u_{kN}\right), \tag{4.8}$$

which is a subset of the network spectrum assignment vector given in (4.6) for the centralized approach.

Algorithms proposed in Chapter 5 are designed so that they can scale from a single cell to a generic number of cells $K$.

#### 4.3.2.3 Cell Characterization Entity

The Cell Characterization Entity has an analogous function to the Network Characterization Entity in the centralized approach, that is, to support the execution of the DSA algorithm by providing an estimation of the performance of the cell under a certain chunk assignment. However, in the decentralized approach, each cell should approximate the spectrum usage of adjacent cells in order to estimate, first the potential intercell interference and then its own chunks' capacities and cell performance.

Two ways of obtaining such information are proposed: 1) a *cooperative spectrum assignment* scheme where cells exchange the current spectrum assignment through an intercell interface (e.g., like the X2 interface in LTE [14]), and 2) a *non-cooperative spectrum assignment* scheme where the spectrum usage of adjacent cells is estimated from users' measurements reports. It is worth to remark that the cooperative spectrum assignment scheme is only possible when a signaling interface between cells is available (e.g.,

[9] In case of simultaneous execution, this could be easily detected e.g., due to consecutive performance degradation in the cell, and then simply another initiation time for the cell would be randomly selected.

macrocell scenarios). On the other hand, the non-cooperative spectrum assignment scheme is possible in most of the scenarios as long as there are measurements reports from users. This is especially useful for femtocell scenarios, where the random location of femtocells makes difficult the deployment of an intercell signaling interface.

Therefore, Cell Characterization Entity could optionally has an extra input in case of cooperative spectrum assignment scheme, which is the spectrum assignment of adjacent cells retrieved from the intercell signaling interface as depicted in Figure 4.5. Then, model implementations given at the end of Chapter 5 will consider both the cooperative and non-cooperative schemes.

#### 4.3.2.4 Execution entity

The execution entity has the same function as the homonymous entity in the centralized approach but at cell level only.

## 4.4 Summary

This chapter has introduced the framework for dynamic spectrum assignment in the ambit of downlink OFDMA cellular networks. One of the main benefits of the proposed framework is that it decouples the resource assignment in two temporal scales, one in the medium-term, that assigns OFDMA frequency chunks to cells, and other that operates in the short-term to assign to each user the previously assigned frequency chunks to a given cell. This approach decreases the complexity of the frequency resource assignment in terms of both signaling and computational load.

Moreover, two versions of the framework have been presented. A centralized approach with a central entity that decides the spectrum assignment for a set of cells. This approach benefits from handling global information so that interactions between cells (i.e., intercell interference) could be easily detected. Also, thanks to the medium-term execution, it has been proven that the signaling exchange between cells and the central entity is not a limitation. However, the approach could not be practical for scenarios where this central control entity is not available, such as in emergent femtocell networks.

Therefore, a decentralized approach has been proposed as an evolution of the centralized one. In the decentralized framework, each cell autonomously and independently decides its spectrum assignment, but considering the spectrum usage of the cells in the considered cell's neighborhood. Aspects such as the stability of the framework and the cooperation between cells to cope with the spectrum assignment decision have been considered.

Finally, the introduced framework for dynamic spectrum assignment is based in self-organization, so that, observation, analysis, decision, learning and execution functions are included to provide the network/cells with the ability of autonomously changing the spectrum assignment as required. The decision function resides on the so called DSA algorithm. The following chapter is devoted to present two proposals to implement this algorithm based on heuristics and reinforcement learning respectively.

# 5

# DYNAMIC SPECTRUM ASSIGNMENT ALGORITHMS

## Outline

# 5 Dynamic Spectrum Assignment Algorithms

Frequency Reuse Factor schemes described in Chapter 3 deploy a fixed reuse pattern over the network that limits system's performance. These strategies are static and inflexible since the assignment of frequency resources to cells is homogeneous and, in general, cannot be changed online. This implies that the frequency deployment may not be adapted to heterogeneous spatial traffic distributions and their variation in time. Moreover, it is difficult to find a group of cells where the same spectrum band is not used, what prevents them from being offered to a spectrum secondary market.

In this chapter a set of DSA (Dynamic Spectrum Assignment) algorithms is proposed. These algorithms are self-adaptive in the sense that they learn from experience how to better adapt to network conditions, and particularly, how to dynamically adjust the spectrum assignment per cell to temporal and spatial variations of the traffic load. Therefore, in order to avoid wasting frequency resources and to generate spectrum pools for secondary usage, different cells are given different number of chunks depending on their traffic load, channel conditions and users' QoS requirements.

DSA algorithms are part of the DSA framework proposed in Chapter 4. There, the Network/Cell DSA controller (in charge of providing the network/cell spectrum assignment in the medium term for the centralized/decentralized approaches respectively) executes a DSA algorithm whose mission is to deliver a spectrum assignment vector as an output for the network/cell. That is, the objective of the DSA algorithm is to build an assignment vector $\Upsilon = (\Upsilon_1, \Upsilon_2, ..., \Upsilon_K)$ where $\Upsilon_k = (u_{k1}, ..., u_{kn}, ..., u_{kN})$ is the $1 \times N$ chunk-to-cell assignment vector for cell $k$ out of $K$ cells. Concretely, if the $n$-th chunk has to be assigned to the $k$-th cell then $u_{kn}$ =1, otherwise, $u_{kn}$ =0.

In order to support DSA algorithm execution, the Network/Cell characterization entity in the framework gives an estimation of the response of the network/cell to a given spectrum assignment. Moreover, the Status Observer module provides to both, the DSA algorithm and the Network/Cell characterization entity an execution context, which is compound of the average number of users per cell denoted as $U_k$, the average dissatisfaction probability per cell $P_k^{th_{\text{target}}}$, and, in case of the decentralized approach of the framework, the probability density function of the average SINR for each chunk ($\overline{\gamma}_n$), denoted as $s_{\overline{\gamma}_n}(\overline{\gamma}_n)$.

Two methodologies have been considered as foundations of the DSA algorithms. First, section 5.1 presents a set of DSA algorithms based on heuristics (called HEUR-DSA), which are intuitive and simple forms of solving the DSA problem. However, they are sub-optimal. Then, a Reinforcement Learning algorithm for DSA (RL-DSA) is presented in section 5.2. The inherent optimization behavior of RL over a reward signal is exploited by RL-DSA to dynamically find proper spectrum assignments per cell depending on current traffic load variations, at the same time that *learning* is retained to be exploited in subsequent spectrum assignment tasks.

Finally, section 5.3 gives a description of the implementation of the network/cell characterization entity included in the DSA framework (see Chapter 4 subsections 4.2.2.3 and 4.3.2.3), since its concrete implementation could depend on the DSA algorithm, (for instance, in case of RL-DSA algorithm, it would be in charge of building the reward signal).

## 5.1 Heuristic algorithm

Heuristic methodology has been used to provide a reasonable solution to the spectrum assignment in a cellular scenario by following an intuitive procedure that is compound of two phases. The high level description of the procedure executed by the algorithms is schematically represented in Figure 5.1. First, the number of chunks to assign to each cell is computed considering cells' traffic loads and users' QoS requirements. Afterwards, an assignment procedure that mitigates intercell interference (cost of assignment) is executed to decide the specific chunks to be assigned per cell. It is worth to mention that, as in general for heuristic algorithms, the solution given by the HEUR -DSA algorithm is not proven to be optimal, but performance results reveal its effectiveness in practical scenarios, as it will be shown in Chapter 7.

Four variants of the HEUR -DSA algorithm named *HEUR-DSA1*, *HEUR-DSA2*, *HEUR-DSA3* and *HEUR-DSA4* are introduced to cope with the limitations of the FRFs.

HEUR-DSA1 and HEUR-DSA2 algorithms follow the same deployment as FRF1 and FRF3 (i.e., the same power is devoted to all chunks and there is no division between central and edge sets, neither in the spectrum nor in the cell and users, see Chapter 3 subsection 3.1.2 for details). The HEUR-DSA1 algorithm estimates the number of chunks per cell based on the number of users per cell and their throughput requirements. However, the HEUR-DSA1 algorithm considers a fixed chunk capacity to estimate the number of chunks needed. In consequence, it may not properly adapt the number of chunks per cell if the real average capacity per chunk differs from the original estimation. The HEUR-DSA2 algorithm self-tunes its parameters to better adjust to the real average capacity per chunk and consequently better adapts the number of chunks to traffic requirements per cell. Both algorithms execute the same chunk-to-cell assignment procedure (phase 2) that is based on minimizing the cost assignment per chunk and cell.

HEUR-DSA3 and HEUR-DSA4 follow the same approach as Partial frequency Reuse (PR) and Soft frequency Reuse (SR) respectively by dividing the spectrum, cells, and users between central and edge subsets. Also the power per chunk depends on its usage (i.e., the power devoted to edge usage is greater than that to central usage, see also 3.1.2 for details). The key of this implementation is to protect the users at the edge of the cell, providing them a better SINR. As the HEUR-DSA2 algorithm, the HEUR-DSA3 and HEUR-DSA4 algorithms also implement a self-tuning mechanism.

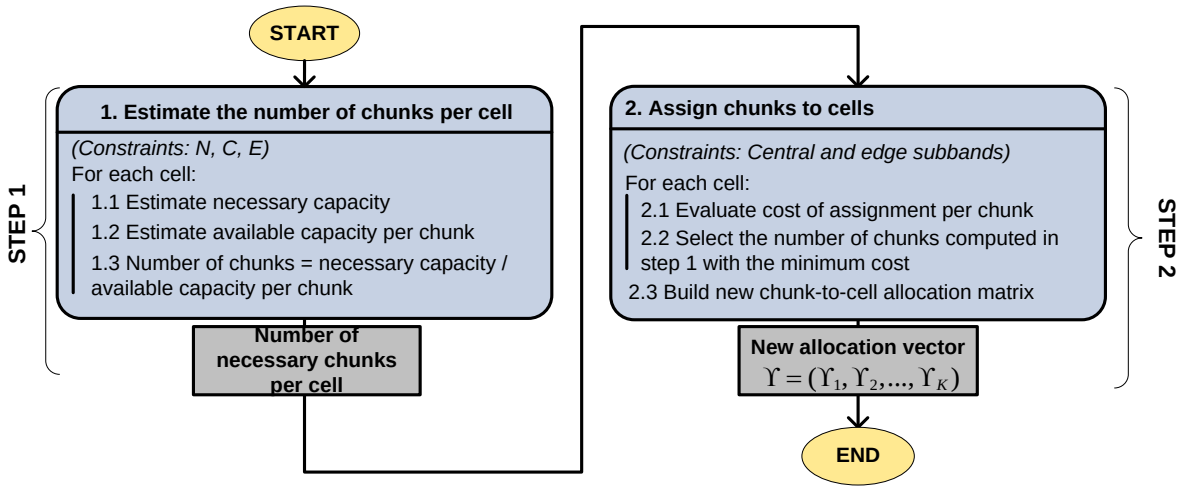


**Figure 5.1 Steps of the HEUR-DSA algorithm**

Following sections are devoted to describe the implementation details for each one of the variants of the HEUR-DSA algorithm.

## 5.1.1 HEUR-DSA functional description

### 5.1.1.1 HEUR-DSA1 Algorithm

The HEUR-DSA1 algorithm executes the following actions in each phase:

*Phase 1*: *Compute the number of chunks to assign in each cell*. The number of chunks is adapted to cells' traffic load (i.e., number of users) so that highly loaded cells get a high number of chunks. Specifically, given a maximum number of chunks $N$ available in the system, the number of chunks $N_k \in \{1,2,...,N\}$ assigned to the $k$-th cell is given by

$$N_k = \min\left(N, \max\left(1, \left\lceil \frac{U_k th_{\text{target}}}{(W/N)(\eta_{\max}/\varpi)} \right\rceil\right)\right), \tag{5.1}$$

where $\lceil x \rceil$ denotes the nearest integer greater than or equal to $x$ and $U_k$ are the users served by $k$-th cell. This value is included in the execution context provided by the Status Observer in the DSA framework presented in Chapter 4. $th_{\text{target}}$ represents the satisfaction throughput per user (i.e., the minimum throughput that a user expects to be satisfied with the requested service), and $(W/N)(\eta_{\max}/f)$ is the estimated chunk capacity in bits/s where $W$ is the total system bandwidth (and correspondingly $W/N$ is the chunk bandwidth), $\eta_{max}$ stands for the maximum theoretical spectral efficiency due to AMC in bits/s/Hz, and $\varpi \geq 1$ is an empirical *margin factor*. It is expected that due to poor channel conditions (especially for users at the edge of the cell) the average spectral efficiency obtained in the cell is lower than $\eta_{max}$. Then, the margin factor reduces $\eta_{max}$ in order to obtain a closer value to the actual chunk capacity. For HEUR-DSA1 $\varpi$ is fixed and configured offline. Hence, the chunk capacity might be over- or under-estimated depending on the value of $\varpi$.

*Phase 2*: *Assign the chunks devoted to each cell determining the potential intercell interference*. The HEUR-DSA1 algorithm assigns to a given cell $k$ the necessary chunks ($N_k$ from the step 1) with the minimum cost. The cost of assigning the $n$-th chunk to cell $k$ is denoted in the following as $\vartheta_k^{(n)}$, and is computed by the Network/Cell Characterization Entity of the DSA framework as detailed in section 5.3. Roughly it takes into account the potential intercell interference between cells, and the load of the cells.

A detailed description of the phase 2 of the HEUR-DSA1 algorithm is included in Table 5.1.

### 5.1.1.2 HEUR-DSA2 Algorithm

Unlike the HEUR-DSA1 algorithm, where the margin factor $\varpi$ remains constant, the HEUR-DSA2 algorithm adjusts the margin factor per cell $\varpi_k$ according to the average dissatisfaction probability on the cell $k$ ($P_k^{th_{\text{target}}}$). Thus, it provides a better estimation of the capacity per chunk than HEUR-DSA1, through a self-tuning mechanism. The margin factor $\varpi_k$ is updated as:

$$\varpi_k = \begin{cases} \varpi_k - \Delta\varpi, & if\ P_k^{th_{\text{target}}} \le P_{down} \\ \varpi_k, & if\ P_{down} \le P_k^{th_{\text{target}}} \le P_{up} \\ \varpi_k + \Delta\varpi, & if\ P_k^{th_{\text{target}}} \ge P_{up} \end{cases} \tag{5.2}$$

where $\Delta\varpi$ is defined as the *margin factor step*, and $P_{down}$ and $P_{up}$ are dissatisfaction probability thresholds to decrease and increase the margin factor respectively.

*Phase 1*: *Compute the number of chunks to assign in each cell.* After updating the margin factor using (5.2), the number of chunks to assign to cell $k$ is then computed using (5.1) after substituting $\varpi$ by $\varpi_k$. Notice that the HEUR-DSA2 tends to maintain the dissatisfaction probability between $P_{down}$ and $P_{up}$ if possible, which reduces the number of chunks used per cell if dissatisfaction probability is below $P_{down}$. This behavior improves the spectral efficiency and increases the number of free resources for secondary usage.

*Phase 2*: *Assign the chunks devoted to each cell determining the potential intercell interference.* The HEUR-DSA2 algorithm implements the same assignment procedure as HEUR-DSA1 (Table 5.1).

#### 5.1.1.3 HEUR-DSA3 Algorithm

Similar to the Partial frequency Reuse (PR) strategy, the HEUR-DSA3 algorithm considers that the system's band is divided into two separate bands with available chunks devoted to central and edge usage respectively. The maximum number of chunks per subband is $C$ and $E$ for the central and the edge subbands respectively so that $N = C + E$. This division is configured by the network operator and remains unaltered for all cells so that central chunks never perceive interference from edge chunks and *vice versa*.

The HEUR-DSA3 algorithm (a) computes the number of chunks needed per subband and per cell (step 1) and (b) performs the chunk-to-cell assignment independently per subband (step 2). As in the PR strategy, no intercell mitigation is performed for the central subband whereas for the edge subband the same assignment strategy as in HEUR-DSA1 and HEUR-DSA2 algorithms is executed. Thus, HEUR-DSA3 is a dynamic version of PR.

*Phase 1: Compute the number of chunks to assign in each cell.* HEUR-DSA3 algorithm computes the number of chunks to assign to the center band $C_k$ and the edge band $E_k$ for the $k$-th cell as

$$C_k = \min\left(C, \max\left(1, \left\lceil \frac{U_{Ck} th_{\text{target}}}{(W/N)(\eta_{\max}/\varpi_k)} \right\rceil\right)\right), \tag{5.3}$$

$$E_k = \min\left(E, \max\left(1, \left\lceil \frac{U_{Ek} th_{\text{target}}}{(W/N)(\eta_{\max}/\varpi_k)} \right\rceil\right)\right), \tag{5.4}$$

where $U_{Ck}$ and $U_{Ek}$ are the number of users located in the central and edge part of the $k$-th cell respectively and the margin factor per cell $\varpi_k$ is updated using (5.2). Notice that the number of chunks per subband is restricted to the same $C$ and $E$ for any cell what could be a limitation if the cells have different loads. On the other hand, this process eases the intercell interference mitigation process, especially for edge chunks as is explained next.

**Table 5.1 HEUR-DSA spectrum assignment procedure (PHASE 2)**

**Definitions**

$K$ : number of cells considered.
$k \in \{1,2,...,K\}$ : cell index.
$N$ : number of chunks available.
$N_k$ : number of chunks to assign to cell $k$
$C$ : maximum number of chunks in central subband
$C_k$ : number of central chunks to assign to cell $k$
$E$ : maximum number of chunks in edge subband
$E_k$ : number of edge chunks to assign to cell $k$
$n \in \{1,2,...,N\}$ : chunk index.
$\Upsilon_k$ : 1xN assignment vector for cell $k$.
$\Upsilon = (\Upsilon_1, \Upsilon_2, ..., \Upsilon_K)$ : Network assignment vector
$\Phi_n$ : set of cells with chunk $n$ assigned.
$\Phi_n^{(E)}$ : set of cells with chunk $n$ assigned for edge usage.
$\vartheta_k^{(n)}$ : cost of assigning chunk $n$ to cell $k$.
$\vartheta_k^{(n)}\big|_{edge}$ : cost of assigning chunk $n$ to cell $k$ for edge usage.
$\mathbf{0}_N$ : null 1xN vector

**HEUR-DSA1 and HEUR-DSA2**

**1:** **Init-** $\Phi_n = \varnothing \ \forall n$,
$\Upsilon_k = \mathbf{0}_N \ \forall k$,
$\vartheta_k^{(n)} = 0 \ \forall n,k$
**2:** **for** all $k$, **do**
Compute costs:
**3:** **for** all $n$, **do**
Obtain cost $\vartheta_k^{(n)}$
**4:** **end for**
Perform assignment:
**5:** **while** $\sum_{n=1}^{N} \Upsilon_k[n] < N_k$, **do**
$n^* = \underbrace{\arg\min}_{\forall n} \left\{ \vartheta_k^{(n)} \right\}$
$\Upsilon_k[n^*] = 1$
$\Phi_{n^*} \leftarrow \Phi_{n^*} \cup \{k\}$
$\vartheta_k^{(n^*)} = \infty$
**6:** **end while**
**7:** Update Assignment vector:
$\Upsilon = (\Upsilon_1, \Upsilon_2, ..., \Upsilon_K)$
**8:** **end for**

**HEUR-DSA3**

**1:** **Init-** $\Phi_n^{(E)} = \varnothing \ \forall n$,
$\Upsilon_k = \mathbf{0}_N \ \forall k$,
$\vartheta_k^{(n)} = 0 \ \forall n,k$
**2:** **for** all $k$, **do**
Reuse 1 for central subband:
**3:** $n \leftarrow 1$
**4:** **while** $n \le C_k$, **do**
$\Upsilon_k[n] = 1$
$n \leftarrow n+1$
**5:** **end while**
Compute cost edge subband:
**6:** **for** all $n$ >C, **do**
Obtain cost $\vartheta_k^{(n)}\big|_{edge}$
**7:** **end for**
**8:** Assignment edge subband:
**9:** **while** $\sum_{n=1}^{N} \Upsilon_k[n] < C_k + E_k$, **do**
$n^* = \underbrace{\arg\min}_{\forall n>C} \left\{ \vartheta_k^{(n)}\big|_{edge} \right\}$
$\Upsilon_k[n^*] = 1$
$\Phi_{n^*}^{(E)} \leftarrow \Phi_{n^*}^{(E)} \cup \{k\}$
$\vartheta_k^{(n)}\big|_{edge} = \infty$
**10:** **end while**
**11:** Update Assignment vector:
$\Upsilon = (\Upsilon_1, \Upsilon_2, ..., \Upsilon_K)$
**12:** **end for**

**HEUR-DSA4**

**1:** **Init-** $\Phi_n^{(E)} = \varnothing \ \forall n$,
$\Upsilon_k = \mathbf{0}_N \ \forall k$,
$\vartheta_k^{(n)} = 0 \ \forall n,k$
**2:** **for** all $k$, **do**
Compute cost edge subband:
**3:** **for** all $n$, **do**
Obtain cost $\vartheta_k^{(n)}\big|_{edge}$
**4:** **end for**
Assignment edge subband:
**5:** **while** $\sum_{n=1}^{N} \Upsilon_k[n] < E_k$, **do**
$n^* = \underbrace{\arg\min}_{\forall n} \left\{ \vartheta_k^{(n)}\big|_{edge} \right\}$
$\Upsilon_k[n^*] = 1$
$\Phi_{n^*}^{(E)} \leftarrow \Phi_{n^*}^{(E)} \cup \{k\}$
$\vartheta_k^{(n)}\big|_{edge} = \infty$
**6:** **end while**
Compute cost central subband:
**7:** **for** all $n$, **do**
**8:** **if** $\Upsilon_k[n] = 0$, **then**
Obtain cost $\vartheta_k^{(n)}\big|_{edge}$
**9:** **end if**
**10:** **end for**
**11:** **while** $\sum_{n=1}^{N} \Upsilon_k[n] < C_k + E_k$, **do**
$n^* = \underbrace{\arg\min}_{\forall n} \left\{ \vartheta_k^{(n)}\big|_{edge} \right\}$
$\Upsilon_k[n^*] = 1$
$\vartheta_k^{(n)}\big|_{edge} = \infty$
**12:** **end while**
**13:** Update Assignment vector:
$\Upsilon = (\Upsilon_1, \Upsilon_2, ..., \Upsilon_K)$
**14: end for**

*Phase 2*: *Assign the chunks devoted to each cell determining the potential intercell interference.* The assignment algorithm is independently executed for each subband. Chunks devoted to the central subband are assigned to the cells regardless the intercell interference because the intercell interference for these chunks is reduced due to the lower power that they are transmitted. Then, in this step, central chunks are consecutively assigned to a cell from the beginning of the central subband. On the other hand, the same assignment procedure as in HEUR-DSA1 and HEUR-DSA2 algorithms is executed for the edge subband. Therefore, the HEUR-DSA3 algorithm conserves the same assignment policy as PR but applying a dynamic spectrum assignment in the edge subband. The assignment method of the HEUR-DSA3 algorithm is detailed in Table 5.1.

#### 5.1.1.4 HEUR-DSA4 Algorithm

The HEUR-DSA4 algorithm goes one step further and automatically adapts the number of chunks dedicated in each cell to the central and edge subbands without operator's intervention. It also makes a division between central and edge chunks, adapting their number to central and edge users' requirements respectively (step 1). Next, as in the SR scheme, the HEUR-DSA4 approach permits to assign the central and edge chunks at any place within the system's band. Thus, the chunk assignment for each subband is not restricted to a specific area of the spectrum and, for example, a chunk could be assigned for edge usage in a cell even if the same chunk is reserved to central usage in other cells. HEUR-DSA4 constitutes a dynamic version of SR.

*Phase 1: Compute the number of chunks to assign in each cell.* The HEUR-DSA4 algorithm computes the number of chunks to assign to the center band $C_k$ and the edge band $E_k$ for the $k$-th cell as

$$C_k = \max\left(1, \left\lceil \frac{U_{Ck} th_{\text{target}}}{(W/N)(\eta_{\max}/\varpi_k)} \right\rceil\right), \tag{5.5}$$

$$E_k = \max\left(1, \left\lceil \frac{U_{Ek} th_{\text{target}}}{(W/N)(\eta_{\max}/\varpi_k)} \right\rceil\right), \tag{5.6}$$

where the total number of chunks in the cell is $N_k = C_k + E_k$. In case that the resulting number of chunks $N_k$ is greater than the maximum available chunks $N$ a further adjustment is carried out. Specifically,

$$C_k \leftarrow \left[ N \frac{C_k}{N_k} \right] \tag{5.7}$$

and

$$E_k \leftarrow \left[ N \frac{E_k}{N_k} \right], \tag{5.8}$$

where $[x]$ denotes the nearest integer to $x$. If still $N_k \neq N$ then one chunk is added to the subband with fewer chunks or subtracted from the subband with more chunks depending on whether $N_k < N$ or $N_k > N$ respectively.

*Phase 2*: *Assign the chunks devoted to each cell determining the potential intercell interference.* As in the SR scheme, the HEUR-DSA4 approach permits to assign the chunks for the center and edge subbands

at any place within the system's band. In any case, the algorithm tries, on the one hand, to assign the best chunks for edge usage and, on the other hand, to minimize the interference generated in chunks assigned for edge usage in previous cells. First, the edge chunks are assigned in order to assure the best combination to the edge subband and next, the central chunks are assigned. In both cases, the cost of assigning the $n$-th chunk to cell $k$ is estimated as $\vartheta_k^{(n)}\big|_{edge}$, where it is considered only the contribution from cells with chunk $n$ assigned for edge usage.

Notice that assigning a specific chunk $n$ when is reserved in the neighboring cells for central usage has no cost. The reason to that is that if the chunk $n$ is going to be used for central usage in two cells, then the reduced radius of the inner cell assures intercell interference protection for central users. On the other hand, if the chunk $n$ is going to be used for edge usage in one cell and for central usage in other cell then the assignment does not constitute a risk if a cautious setting of the powers devoted to central and edge chunks is done. As for the other variants, the assignment procedure for the HEUR-DSA4 algorithm is described in Table 5.1.

Finally, it is worth to remark that the procedure given in Table 5.1 is for the centralized implementation of the algorithm in the Network DSA controller. However, decentralized version is straightforward by simply executing the steps within the 'for' loop in line 2 independently in each cell in the Cell DSA controller. Furthermore, the cost computation is given in section 5.3.

### 5.1.2 Validation

This section first validates proposed HEUR-DSA algorithm in its four variants by illustrating how it adapts the number of chunks per cell depending on the traffic load, and thus, revealing its adaptability. Nevertheless, extensive performance comparison results will be detailed in Chapter 7.

Dynamic simulations for a downlink OFDMA multicell scenario are carried out with the simulation tool that will be explained in Chapter 6. The cellular scenario is composed of $K$=19 omnidirectional cells. The maximum number of chunks in the system is $N$=12 that is big enough to provide frequency diversity. Chunk bandwidth is 375 kHz and maximum spectral efficiency is $\eta_{\max}$ = 4 bits/s/Hz. Hexagonal cells of radius $R$=0.5 km are employed. Users are uniformly distributed within a cell and their mobility is restricted to the cell where they belong to in order to maintain the load ratio between cells, (i.e., no handover effects are modeled).

Users are deployed in the scenario with their buffers always full and all users request a satisfaction throughput $th_{\text{target}}$=128Kbps. This means that a user has always information to transmit and then, he/she aims to get as much capacity as possible above 128Kbps. If available, edge users' transmissions have priority over edge chunks. Then the scenarios are under the worst traffic load given a number of users in the scenario, which is equal to 300 users. HEUR-DSA is configured with $\varpi$=2.5, $\Delta\varpi$=0.05, $P_{low}$ = 0.1%, $P_{up}$ = 5%, $C$=3, $E$=9. Two types of cells are of interest in this study: a cell that progressively increases its traffic load, and alternatively, another cell that progressively decreases its load. It is expected that HEUR-DSA algorithm varies the number of chunks per cell accordingly to the traffic load variation per cell.

Figure 5.2 depicts the average number of chunks assigned to the increasing and decreasing traffic load cells (represented as the load percentage over the load in the whole scenario, i.e., 100% is equal to 300 users), where it is appreciated that HEUR-DSA algorithm increases the number of chunks to cope with the increasing traffic demand. On the other hand, HEUR-DSA decreases the number of chunks per cell if the traffic load demand decreases as well. It will be shown in Chapter 7 that this fact translates into a potential reduction of the intercell interference and also in an increment of the free spectrum for underlay secondary spectrum usage. Notice, however, the inflexibility of the FRFs (i.e., FRF1, FRF3, PR and SR), which do not vary their number of assigned chunks per cell for the variations of the traffic load. Hence, for the increasing load cell there is a lack of capacity for high traffic load percentages.

Finally, comparing HEUR-DSA variants, HEUR-DSA2 demonstrates the best adaptability by using the lowest number of chunks per cell and traffic load in practically all cases. This is because, compared with HEUR-DSA1, HEUR-DSA2 varies the margin factor obtaining a more precise estimation of the capacity per chunk. On the other hand, compared with HEUR-DSA3 and HEUR-DSA4, HEUR-DSA2 has a higher degree of freedom to select the chunks (the whole band), without having to reserve chunks for the central or edge subbands.

## 5.2 Reinforcement Learning algorithm

Chapter 2 described Reinforcement Learning (RL) as a methodology that is able to learn the best actions to apply in order to obtain the best reward or payoff from a variable environment after a large number of interactions with it. Three main benefits can be stated then for RL: (*i*) It naturally optimizes a given reward signal, (*ii*) the optimization objective can be easily changed by modifying the reward signal without varying the RL optimization procedure, and (*iii*) it can retain learning acquired in a given execution to be exploited in subsequent optimization tasks.

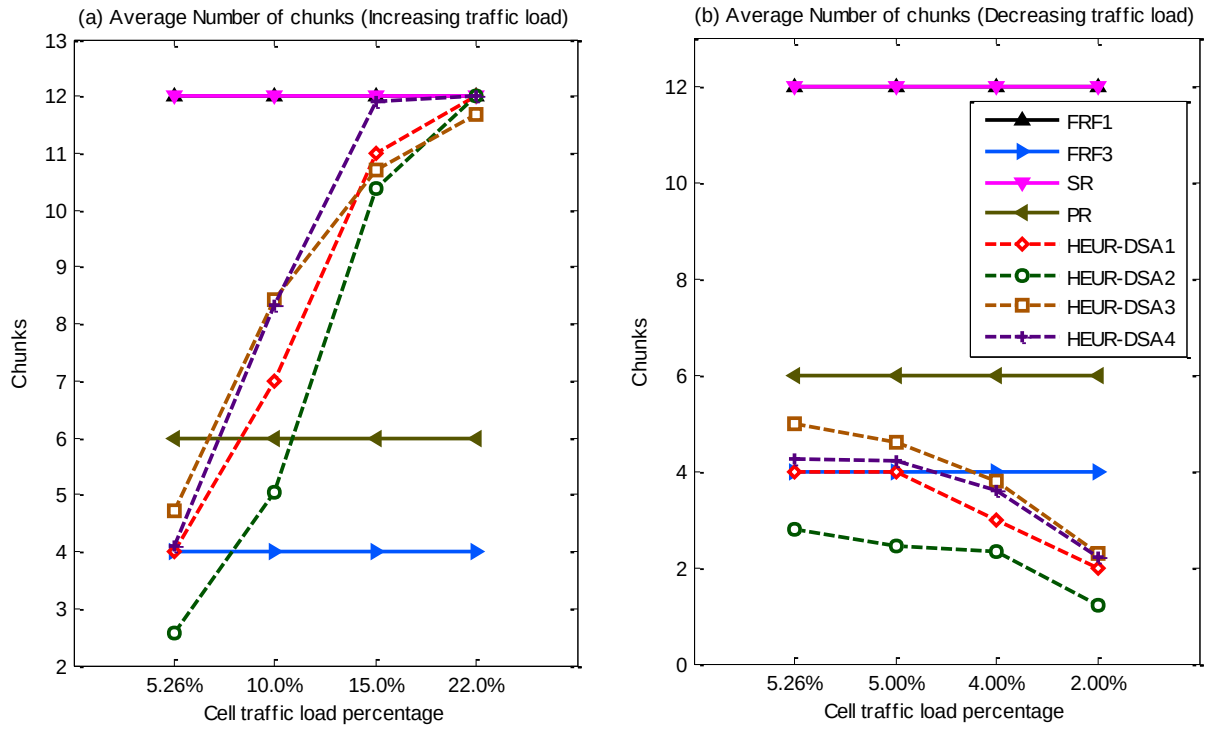


**Figure 5.2 Adaptability of the HEUR-DSA algorithms**

This section presents a RL-based DSA algorithm (RL-DSA) that tries to exploit the abovementioned features in the context of spectrum assignment in OFDMA cellular networks. Each action of RL-DSA represents a candidate spectrum assignment for the network/cell depending on whether the centralized or decentralized approaches of the DSA framework are employed respectively. In turn, the network/cell characterization entity returns a *reward* signal representing the suitability of a given action. This reward signal allows RL-DSA to *learn* the most appropriate spectrum assignment from the point of view of reward maximization.

## 5.2.1 RL-DSA functional description

### 5.2.1.1 REINFORCE methods overview

RL-DSA foundations reside on the REINFORCE RL methods detailed in subsection 2.3.3 in Chapter 2 and summarized here. It has been shown [29] that these methods provide global optimization of a reward signal for a sufficiently large number of interactions with an environment in a succession of RL steps $t$. The building block of RL-DSA is the REINFORCE *Bernoulli Logistic Unit* (BLU) depicted in Figure 2.10. For a generic agent $i$, the output is a two-action Bernoulli random variable $y_i(t) \in Y = \{0,1\}$ with a given action selection probability $p_i(t)$ .That is,

$$y_i(t) / \begin{cases} \Pr(y_i(t) = 1) = p_i(t); \\ \Pr(y_i(t) = 0) = 1 - p_i(t). \end{cases} \tag{5.9}$$

The internal probability depends on the internal status of the agent $w_i(t)$ and the agent's input $x_i(t)$ by means of the logistic function as given in (2.22). Then, for each action $y_i(t)$ the environment returns a given reward $r_i(t)$ that is used by the agent to update its internal status and learn according to the following learning rule:

$$w_i(t) = w_i(t-1) + \Delta w_i(t)\,, \tag{5.10}$$

$$\Delta w_i(t) = \alpha(t)\big(r_i(t) - \overline{r}_i(t-1)\big)\big(y_i(t-1) - p_i(t-1)\big)x_i(t-1) + \alpha(t)\xi\big(w_i(t-1)\big) + \sqrt{\alpha(t)}\zeta_i(t)\,, \tag{5.11}$$

where this rule has been derived from (2.23), (2.24) and (2.26). Details regarding the terms in this rule are given in subsection 2.3.3.2 in Chapter 2. Briefly, $\alpha$ is a learning rate, $\overline{r}_i$ is an average reward, $\xi\big(w_i\big)$ is a bounding function to assure a minimum exploratory action selection probability $p_{\text{explore}}$, and $\zeta_i$ is a perturbation value that takes a value equal to $+\sigma$ or $-\sigma$ with equal probability to give the algorithm the ability of performing global optimization.

In the following, RL-DSA functional architecture and operation procedure is described.

### 5.2.1.2 RL-DSA functional architecture and procedure

Figure 5.3 depicts the functional architecture of RL-DSA, which is composed of $KN$ RL agents, that is, the $kn$-th agent is devoted to learn whether the $n$-th chunk out of $N$ chunks is assigned to the $k$-th cell out of $K$ cells. The functional architecture depicted is for the algorithm implemented in the centralized version of the DSA framework (i.e., in the Network DSA controller). Nevertheless the

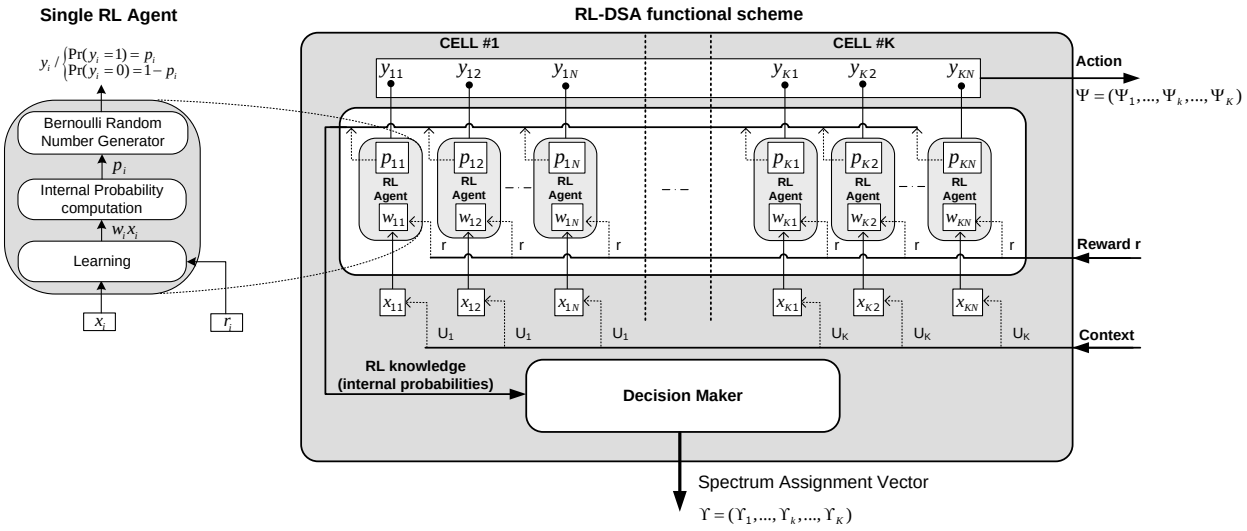


**Figure 5.3 RL-DSA functional architecture**

functional architecture for the decentralized approach of the DSA framework is obtained by simply implementing in the Cell DSA controller an RL-DSA algorithm with only $N$ RL agents. Hence, in the following the general centralized functional architecture is considered, being the decentralized approach a particular case.

In order to face a real world problem with RL it is necessary to appropriately select the physical meaning of the context $x_{kn}$, output $y_{kn}$, and reward $r_{kn}$ for each agent. Particularly, RL-DSA associates different cellular network traffic distributions (context inputs) to different spectrum assignments (output actions). Then, the context $x_{kn}$ reflects the load status of the $k$-th cell. That is, $x_{kn} = U_k \ \forall n$, being $U_k$ the average number of users in the $k$-th cell. This context remains constant during a RL-DSA execution so that RL-DSA is able to associate solutions to both homogeneous and heterogeneous spatial distributions of the traffic load (i.e., users per cell).

On the other hand, the action taken by RL-DSA is a binary vector

$$\Psi(t) = (\Psi_1(t), ..., \Psi_2(t), ..., \Psi_K(t)), \tag{5.12}$$

where $\Psi_k(t) = (y_{k1}(t), y_{k2}(t), ..., y_{kn}(t), ..., y_{kN}(t))$ represents a *candidate* chunk-to-cell assignment in each RL step $t$. To this end, it is considered that the $n$-th chunk is assigned to the $k$-th cell if the output $y_{kn}(t)$ is 1 (and not assigned in case $y_{kn}(t)$ is 0). Notice that, $\Psi(t)$ is not the final spectrum assignment vector delivered by the algorithm $\Upsilon$, but a partial result that RL-DSA iteratively test to learn the best spectrum assignment. Particularly, this action is applied to the Network/Cell Characterization entity, which returns the reward signal $r_{kn}(t)$, whose physical meaning is given in next subsection.

Each time that RL-DSA has to be executed (decided by the Status Observer entity on the DSA framework) it executes the procedure detailed in Table 5.2 for a succession of RL-steps $t = 1, 2, \ldots$. Steps from 1 to 9 execute the learning loop, where for each agent the learning rule is evaluated from the received reward, and also a new action for the RL step is computed. Condition in step 9 finishes the learning loop when a maximum number of steps (MAX_STEPS) is reached. On the other hand, steps from 10 to 16 are executed by *Decision Maker* module in Figure 5.3 to decide the final spectrum assignment for the real network. Notice that RL-DSA bases on the knowledge stored in internal

probabilities $p_{kn}(t)$ and not on the very last random action $\Psi(t)$ to decide the spectrum assignment to the network area. Note that a given chunk can be assigned by the algorithm to more than one cell.

Regarding the first execution, the first time that RL-DSA is triggered, full assignment is set, i.e., $y_{kn}(0)=1$, and accordingly $p_{kn}(0)=1-p_{\text{explore}}$ and $w_{kn}(0)=\Omega=\ln\left((1-p_{\text{explore}})/p_{\text{explore}}\right)\ \forall n,k$ (see subsection 2.3.3.2 in Chapter 2 for details). Moreover $\bar{r}_{kn}(0)=0\ \ \forall n,k$. Notice that, thanks to the exploratory probability, the spectrum assignment chosen by RL-DSA in the following steps can be different from full assignment (because the outputs are Bernoulli random variables). This situation triggers the learning of RL-DSA, causing that internal status and consequently action selection probabilities will evolve according to the learning rule until the end of the learning loop. Moreover, the perturbation term in the learning rule makes that the update of the internal status is different for each chunk in one cell even if the reward and the outputs do not vary in two consecutive steps, allowing RL-DSA to escape from local maxima of the reward signal. Finally, in subsequent triggers, RL-DSA begins from the assignment learnt in the previous run, so that the knowledge acquired until that moment in internal status and probabilities is exploited.

As a final remark, it is assumed that the time taken by RL-DSA to converge to a solution is negligible compared with the medium-term time scale in which RL-DSA is executed by Status Observer (e.g., due to traffic load variations). In fact, convergence studies presented in the validation subsection 5.2.2 in this clause reveal that convergence to a solution with a high degree of proximity to the optimal solution can be obtained even if the learning loop is trunked in a moderate practical number of steps MAX_STEPS.

**Table 5.2 RL-DSA procedure**

| | |
|---|---|
| | **Learning loop: Learns the best action selection probabilities** |
| **1.** | **REPEAT** |
| **2.** | Receive reward $r(t)$ from the environment. |
| **3.** | Update average reward $\bar{r}(t)$. |
| **4.** | **FOR** all $k\in\{1,2,...,K\}$ and all $n\in\{1,2,...,N\}$ |
| **5.** | Update internal status $w_{kn}(t)$ following (5.10) and (2.24) |
| **6.** | Compute internal probabilities $p_{kn}(t)$ |
| **7.** | Generate an action $y_{kn}(t)$ as a Bernoulli random variable with action selection probability $p_{kn}(t)$ |
| **8.** | **END FOR** |
| **9.** | **UNTIL** ( $t > MAX\_STEPS$ ) |
| | **Final decision (Decision Maker): Outputs the assignment vector $\Upsilon_k=(u_{k1},...,u_{kn},...,u_{kN})$ for all $k$** |
| **10.** | **FOR** all $k\in\{1,2,...,K\}$ and all $n\in\{1,2,...,N\}$ |
| **11.** | **IF** $p_{kn}>0.5$ |
| **12.** | Assign the $n$-th chunk to the $k$-th cell, i.e., $u_{kn}=1$ |
| **13.** | **ELSE** |
| **14.** | Do not assign the $n$-th chunk to the $k$-th cell, i.e., $u_{kn}=0$ |
| **15.** | **END IF** |
| **16.** | **END FOR** |

#### 5.2.1.3 Reward signal

The reward signal reflects the suitability of a given action of RL-DSA. In order to learn adequately, it is desirable that RL agents receive the same common reward signal in a given RL step $t$, so that all agents are involved in the same goal-oriented problem [30]. Then, if the decentralized approach of the framework is used, then $r_{kn}(t) = r_k(t) \quad \forall n$, where $r_k \in \mathbb{R}$ is a reward signal for a generic cell $k$. On the other hand, if the centralized approach is used, then $r_{kn}(t) = r(t) \quad \forall k,n$, where $r \in \mathbb{R}$ is a common reward signal that could be obtained as a combination of the particular rewards per cell.

Another property that the reward signal has to verify is that is maximal when certain desired performance metrics are fulfilled for the cell/network. Taking this into account, in the following different formal definitions of the reward per cell $r_k(t)$ and network $r(t)$ are given to cope with different performance optimization objectives.

One possibility is to improve the spectrum usage in a cell. Then the reward signal per cell is defined as

$$r_k(t) = \begin{cases} 0, & \text{if } th_k(t) < th_{\text{target}}; \\ \lambda \hat{\eta}_k(t) + \mu(N - N_k(t)), & \text{otherwise.} \end{cases} \tag{5.13}$$

$th_k(t)$ is the estimated average user throughput for cell $k$ in bits/s, $\hat{\eta}_k(t)$ is the estimated average spectral efficiency in bits/s/Hz, and $N - N_k(t)$ is the number of non-used chunks in that cell (i.e., $N_k(t) = \sum_{n=1}^{N} y_{kn}(t)$ is the number of assigned chunks for a given action $\Psi_k = (y_{k1},...,y_{kN})$ of RL-DSA for the $k$-th cell). $\lambda \geq 0$ and $\mu \geq 0$ are appropriate scaling constants. Then, the reward for a given cell is zero if the average user throughput is below the user satisfaction throughput target $th_{\text{target}}$, so that the reward signal retains the QoS. On the other hand, if the QoS is fulfilled in the $k$-th cell, spectral efficiency in the cell is maximized at the same time that opportunities for secondary spectrum usage in primary non-used spectrum are generated in a Private Commons scenario [6].

Another possibility is to simply maximize cell's SINR assuring a given QoS in terms of a minimum average user throughput. That is, the reward per cell can now be defined as

$$r_k(t) = \begin{cases} 0, & \text{if } th_k(t) < th_{\text{target}}; \\ \hat{\gamma}_k(\text{t}), & \text{otherwise.} \end{cases} \tag{5.14}$$

where $\hat{\gamma}_k(t)$ is the estimated average SINR in the cell $k$.

These variants of the reward signal per cell can be used for the decentralized approach of the DSA framework. Hence, as the reward signal is built by the Cell Characterization Entity for each RL step, this module has to be able to estimate the value of $th_k(t)$, $\hat{\eta}_k(t)$ and $\hat{\gamma}_k(t)$ for each action of RL-DSA. This is given in section 5.3.

On the other side, the common reward signal in case of the centralized version of the DSA framework is defined as

$$r(t) = \sum_{k=1}^{K} r_k(t) + \sum_{j=1}^{s(t)-1} jR\,. \tag{5.15}$$

$r_k(t)$ constitutes the reward signal per cell from either (5.13) or (5.14) as desired, $R$ is an upper bound for all $r_k$, and $s(t)$ stands for the number of cells that fulfill a QoS constraint, $r_k(t) > 0$, as explained in the following. $r_k(t)$ in (5.13) or (5.14) is a positive real value, which, in practice, is upper bounded due to the existence of a maximum achievable spectral efficiency $\eta_{\max}$ and a finite number of available chunks (for $r_k(t)$ given in (5.13)), or an achievable SINR (for $r_k(t)$ given in (5.14)). Then, let $R$ be this upper bound that fulfils $0 \le r_k < R \ \forall k$. The inclusion of the second term in (5.15) assures that $r(t)$ increases monotonically with $s(t)$ as proved in the following.

***Proof:*** *$r(t)$ in (5.15) increases monotonically with the number of cells $s(t)$ fulfilling the QoS constraint.*

The reward signal $r(t)$ from expression (5.15) can be bounded as:

$$\sum_{j=1}^{s(t)-1} jR \le r(t) < Rs(t) + \sum_{j=1}^{s(t)-1} jR \,. \tag{5.16}$$

By substituting the well-known result for the arithmetic sum and operating we can get:

$$0.5R[s^2(t) - s(t)] \le r(t) < 0.5R[s^2(t) + s(t)] \,. \tag{5.17}$$

On the other hand, let assume an increase in the number of cells fulfilling the QoS constraint $s(t)$ to $s'(t) = s(t) + 1$. Then, the corresponding reward $r'(t)$ obtained will be lower bounded as:

$$0.5R[s'^2(t) - s'(t)] = 0.5R[s^2(t) + s(t)] \le r'(t) \tag{5.18}$$

and $r'(t) > r(t)$ follows. This proves that $r(t)$ is a monotonically increasing function with the number of cells $s(t)$ that fulfill the QoS constraint (i.e., $r_k > 0$ ).■

Thus, centralized RL-DSA will tend to select a spectrum assignment that maximizes the reward while at the same time maximizing the number of cells fulfilling the QoS constraint. The reward signal in (5.15) is built by Network Characterization Entity, and since is based on the reward per cell, the same estimation procedures for $th_k(t)$, $\hat{\eta}_k(t)$ and $\hat{\gamma}_k(t)$ can be used.

## 5.2.2 Validation

This subsection is devoted to validate proposed RL-DSA algorithm. To this end, the evolution of internal action selection probabilities and reward is evaluated under a simple but controlled static scenario. Complete performance comparison results will be given in Chapter 7. Moreover, the convergence behavior and sensitivity of the algorithm to different values of its main parameters is examined. Finally, as for the validation of the HEUR-DSA algorithm, simulation tool details can be found on Chapter 6.

### 5.2.2.1 Internal action-selection probabilities and reward evolution

Illustrative results presented here have been obtained over a cellular OFDMA scenario composed of $K$ =3 omnidirectional hexagonal cells of radius $R_c$ =0.5 km. The maximum number of chunks in the

system is $N$=6. Maximum spectral efficiency is $\eta_{\max}$ = 4 bits/s/Hz. Users remain static during simulations and their buffers are always full so their traffic model represents a service that demands as much capacity as possible. They are satisfied if the received throughput during last second is above $th_{\text{target}}$=128 kbps. Two traffic loads are tested: 10 users per cell and 30 users per cell.

Centralized DSA framework approach is employed. Accordingly, reward signal in (5.16) is used with a definition of the reward per cell as in (5.13). The following values were configured: $\lambda$=1 and $\mu$=0 as scaling constants to only consider spectral efficiency. Then, upper bound for the reward signal per cell is $R = \lambda \eta_{\max}$=4. On the other hand, the following values are considered for RL-DSA parameters: learning rate $\alpha$=10, reward averaging factor $\beta$=0.1, perturbation term $\sigma$=0.01, exploratory probability $p_{\text{explore}}$=0.1%, and MAX_STEPS=$10^5$.

Figure 5.4 depicts the chunk to cell spectrum assignment decided by RL-DSA for each one of the traffic loads, where a colored square means that that specific chunk is assigned to a cell. For 10 users per cell only 2 chunks are assigned to each cell whereas for 30 users per cell 4 chunks are given, revealing that RL-DSA adapts the spectrum assignment to traffic load needs. Moreover, notice that the spectrum assignment is performed so that intercell interference between cells is mitigated for all chunks. Hence, in the case of 10 users per cell, the spectrum assignment is 'orthogonal', meaning that chunks assigned to a cell are not used in the other two cells. Besides, in the case of 30 users per cell each assigned chunk to a cell has only interference from one of the other two cells.

Figure 5.5 and Figure 5.6 illustrate the RL learning behavior by showing the internal action selection probabilities evolution and average reward evolution per cell respectively. Notice how the internal probabilities in Figure 5.5 evolve either to $1-p_{\text{explore}}$ or $p_{\text{explore}}$ determining the chunk assignment for each cell when the probability is greater than 0.5. On the other hand, the evolution of the average reward per cell depicted in Figure 5.6 reveals that the reward increases for all cells as RL-DSA learns from the Network characterization entity. Concretely, for the case of 10 users per cell, the average reward per cell evolves until the maximal upper bound ($R$=4), because RL-DSA converge to a spectrum assignment that allows attaining the maximal spectral efficiency $\eta_{\max}$=4bits/s per cell (i.e., there is no intercell interference). Alternatively, for the 30 users case it is not possible to attain maximal spectral efficiency due to intercell interference but reward also shows an increasing tendency.

Finally, it is worth to remark the meaning of low peaks in Figure 5.6. They are produced in the learning loop when RL-DSA selects candidate spectrum assignments (actions) that return low rewards. However, even in these situations RL-DSA still learns since discards solutions that seem to be not suitable.

#### 5.2.2.2 Convergence behavior

The convergence behavior of RL-DSA for different values of its main parameters such as the learning rate ($\alpha$), the maximum number of steps (*MAX_STEPS*), and the random perturbation term ($\sigma$) are studied hereafter. These results show a qualitative behavior that may be useful to setup the RL-DSA.

Figure 5.7 shows Root Mean Square Error (RMSE) between the reward achieved by RL-DSA and the optimal reward in a given scenario. In order to make feasible the computation of the optimal reward, we have set a very particular scenario with 19 cells and 19 chunks, and 5 users per cell, where any spectrum

assignment that gives one different chunk per cell (i.e., no intercell interference) was considered to be optimum, that is, attained the best reward.

Note that an excellent RMSE of 1% can be reached in $10^5$ steps for some of the configured parameters, which denotes a good convergence behavior of RL-DSA since the solution space for the scenario involves $2^{19 \cdot 19}$ ($4.69 \cdot 10^{108}$) different assignments. The study reveals that a high $\alpha$ reduces the number of steps needed to converge to the optimal solution. On the other hand, high values of $\sigma$ perform better for a low number of steps but lower values obtain lower RMSE for a high number of steps. Finally, notice that for a number of steps above $10^6$ the RMSE falls below 2% for all tested values of the parameters, revealing a robust behavior of RL-DSA with respect to the values selected for its parameters.

Finally, RL-DSA requires a small constant number of operations per step (i.e. few additions and products including simple forms of random number computation). Furthermore, memory requirements are low since only few records to store the weights, probabilities, rewards and outputs are needed per RL-agent. These properties make the implementation of the RL-DSA scheme quite feasible.

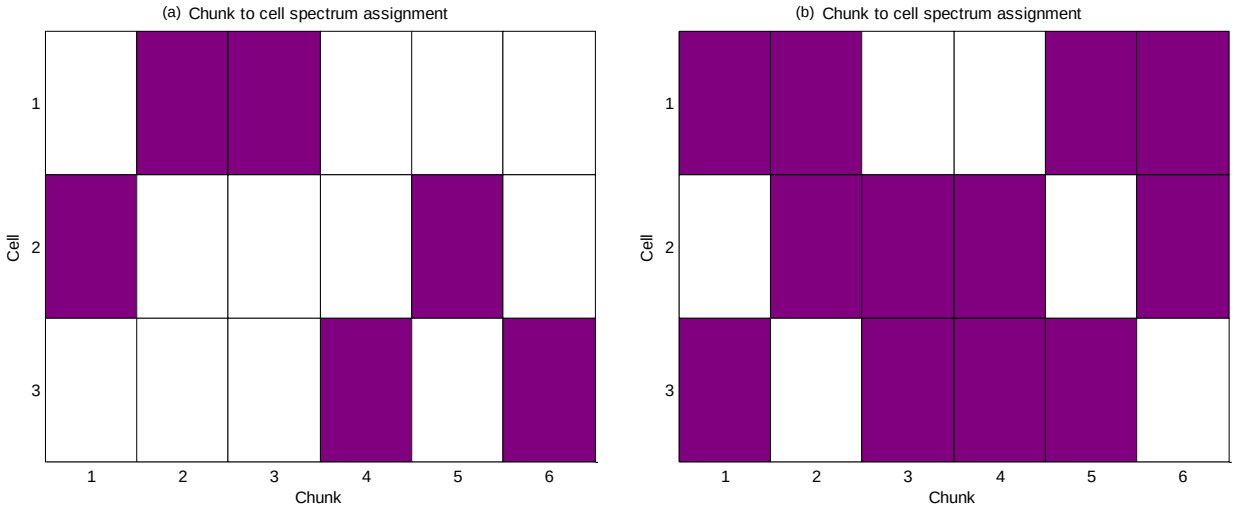


**Figure 5.4 Chunk to cell spectrum assignment. (a) 10 users per cell. (b) 30 users per cell.**

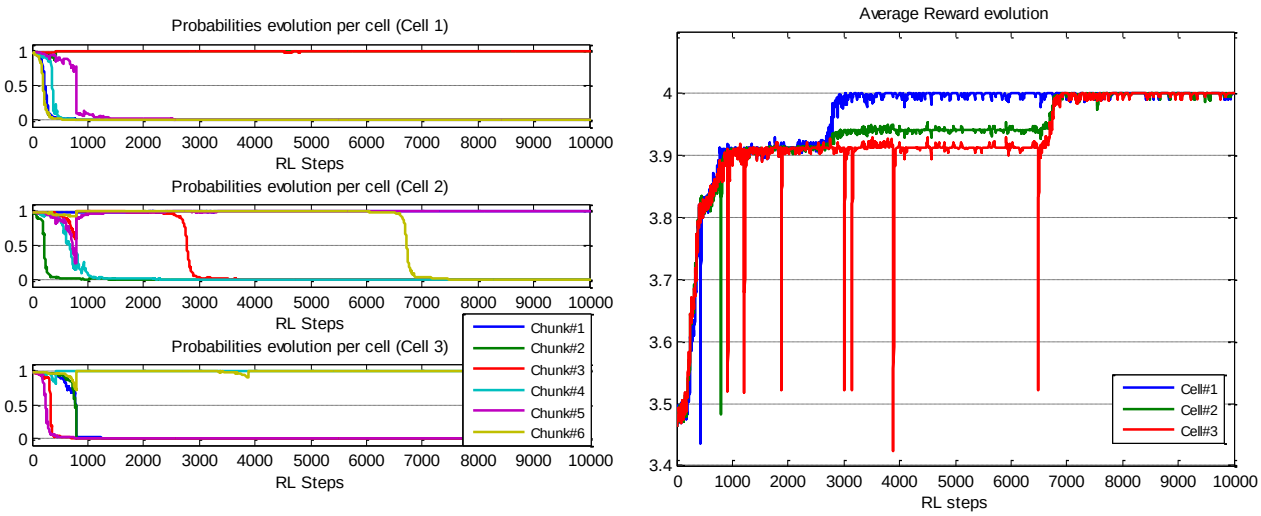


**Figure 5.5 RL internal probabilities and average reward evolution per cell (10 users per cell)**

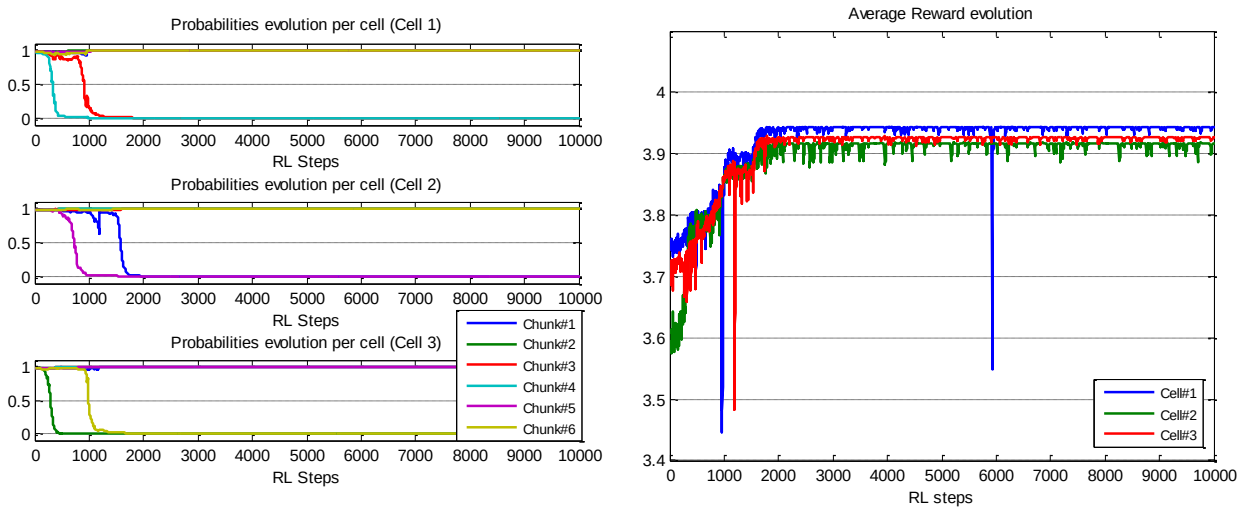


**Figure 5.6 RL internal probabilities and average reward evolution per cell (30 users per cell)**

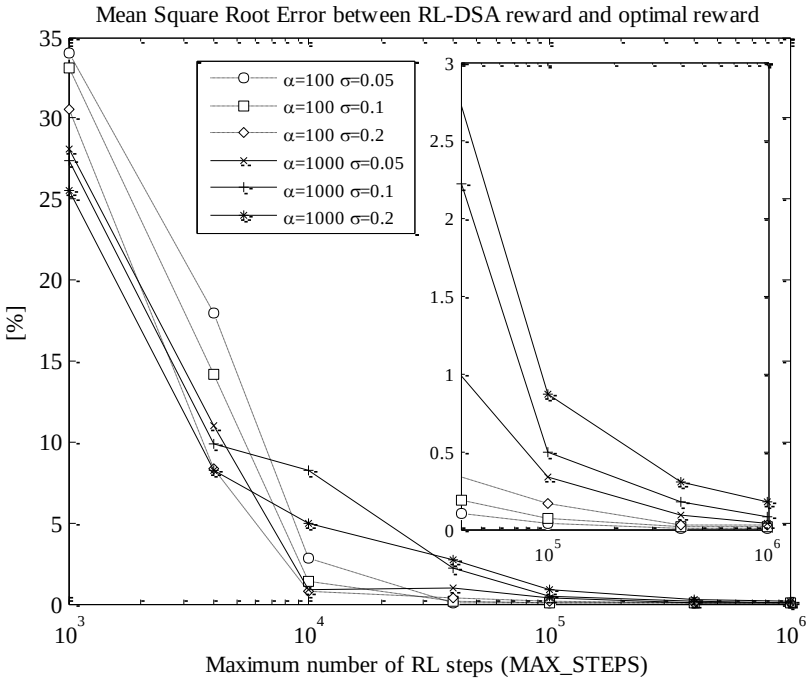


**Figure 5.7 Convergence behavior of RL-DSA**

## 5.3 Network/Cell characterization entity implementation

This section provides a description of the Network/Cell characterization entity in the DSA framework presented in Chapter 4. The aim of this entity is to support the execution of the DSA algorithm by providing an estimation of the performance of the network/cell under a certain spectrum to assignment so that the DSA algorithm can obtain the best chunk-to-cell assignment.

As inputs, Network/Cell characterization entity has the candidate spectrum assignment from the DSA algorithm and the execution context from the Status Observer. As explained in Chapter 4, this execution context includes, the average number of users in a cell, denoted as $U_k$, and the average dissatisfaction

probability per cell $P_k^{th_{\text{target}}}$, and the probability density function of the average SINR for each chunk in each cell ($\bar{\gamma}_n^{(k)}$), denoted as $s_{\bar{\gamma}_n}(\bar{\gamma}_n)$ (index to the cell has been omitted)

Regarding the outputs, depend on the particular implementation and needs of the DSA algorithm. Thus, the Network/Cell characterization entity assists HEUR-DSA algorithm by providing the cost of assignment $\vartheta_k^{(n)}$ and $\vartheta_k^{(n)}\big|_{edge}$ of a given chunk $n$ for a given cell $k$ during the execution of the phase 2 detailed in Table 5.1. Alternatively, this entity returns the reward for the network/cell that RL-DSA employs to learn the best spectrum assignment. More precisely, it has the means to compute estimations of the average user throughput $th_k(t)$, the average spectral efficiency $\hat{\eta}_k(t)$ and the average SINR $\hat{\gamma}_k(t)$ per cell, so that the reward signals given in (5.13), (5.14) and (5.15) can be built for a given RL step $t$.

In the following a proposal to estimate abovementioned values for each one of the proposed DSA algorithms is given.

### 5.3.1 Network/cell characterization for HEUR-DSA

The cost of assigning the $n$-th chunk to an objective cell $k$ is computed as

$$\vartheta_k^{(n)} = \sum_{i \in \Phi_n} \mathbf{A}(i,k), \tag{5.19}$$

where $\Phi_n$ is the set of cells with the $n$-th chunk assigned and $\mathbf{A}$ is a coupling matrix based on the path losses between cells and cells' traffic loads as follows:

$$\mathbf{A}(i,k) = \begin{cases} 0, & \text{if } i = k \\ \left(\frac{U_i}{U_k} + \frac{U_k}{U_i}\right)\left(\frac{R_c}{L_{ik}}\right)^{\chi}, & \text{otherwise} \end{cases} \tag{5.20}$$

where $U_k$ stands for the load of cell $k$ in users, $R_c$ denotes the cell radius, $L_{ik}$ is the minimum distance from center of cell $i$ to the border of cell $k$ and $\chi$ is the path loss propagation exponent. That is, in order to reduce the intercell interference, the cost is given trying to avoid that neighboring cells (high $R_c / L_{ik}$ ratio) use the same chunk. Also, the term $(U_i / U_k + U_k / U_i)$ is included to orient that cells with similar loads reuse the same chunk. The reason is that if one cell is highly loaded with respect their neighbors, cells around it should not re-use the same chunks since this cell possibly would generate interference all the time. Alternatively, if a cell is low loaded then cells around it should avoid using the few chunks assigned to that cell, since it may not benefit from frequency diversity and therefore the interference effects may be worse.

With respect to the cost of assignment of a chunk for edge usage can be computed as

$$\vartheta_k^{(n)}\big|_{edge} = \sum_{i \in \Phi_n^{(E)}} \mathbf{A}(i,k), \tag{5.21}$$

where now $\Phi_n^{(E)}$ is the set of cells with the $n$-th chunk assigned for edge usage.

Notice that in order to compute these costs, it is necessary that the Network/Cell characterization entity knows the chunk assignment for other cells different from the objective cell $k$. This is feasible in the centralized approach of the DSA framework and also for the cooperative decentralized approach, since spectrum assignment is exchanged between adjacent cells. For the non-cooperative decentralized approach, the following cost is proposed based on an estimation of the average SINR per chunk

$$\vartheta_k^{(n)} = \frac{1}{\hat{\gamma}_k^{(n)}}, \tag{5.22}$$

where $\hat{\gamma}_k^{(n)}$ is the average SINR per chunk per cell. It can be computed as

$$\hat{\gamma}_k^{(n)}(t) = \int_{-\infty}^{\infty} \overline{\gamma}_k^{(n)} s_{\overline{\gamma}_k^{(n)}}(\overline{\gamma}_k^{(n)}) d\overline{\gamma}_k^{(n)}, \tag{5.23}$$

from the probability density function of the average SINR for each chunk in each cell $s_{\overline{\gamma}_k^{(n)}}(\overline{\gamma}_k^{(n)})$ provided by Status Observer. With this cost, HEUR-DSA tends to selects those chunks with the better SINR.

Finally, it is worth to remark that HEUR-DSA3 and HEUR-DSA4 cannot be implemented under a non-cooperative decentralized framework because it is not easy or practical to know the usage (central or edge) that other cells are giving to each chunk.

## 5.3.2 Network/cell characterization for RL-DSA

Network/Cell characterization entity has to compute the reward signal for each candidate spectrum assignment provided by RL-DSA each RL step $t$ during the learning loop. Regarding the definitions given for the reward signal in (5.13), (5.14) and (5.15), the Network/Cell characterization entity has to determine estimations of the average user throughput $th_k(t)$, the average spectral efficiency $\hat{\eta}_k(t)$ and the average SINR $\hat{\gamma}_k(t)$ per cell. The way of computing these values depends on whether the centralized or decentralized version of the DSA framework is employed and also on whether exists cooperation between cells or not. The analysis is focused on an objective cell $k$, so index to it will be omitted unless necessary.

### 5.3.2.1 Centralized DSA framework or Cooperative Decentralized DSA framework

In the centralized DSA framework or in the cooperative decentralized DSA framework, Network/Cell characterization entity can know the spectrum assignment of other cells different from the objective cell $k$ (e.g., cells exchange the spectrum assignment in the cooperative decentralized DSA framework). Thus it can compute the set of cells, $\Phi_n(t)$, that cause interference for each chunk $n$ in a given RL step. Then, assuming an interference limited macrocell scenario with omnidirectional antennas, uniformly distributed users per cell, and given a candidate spectrum assignment for the cell in an RL step, the average SINR per chunk in the cell can be estimated as

$$\hat{\gamma}_k^{(n)}(t) = \iint_A \frac{1}{A} SIR(\Phi_n(t), \rho, \theta) \rho d\rho d\theta, \tag{5.24}$$

where $SIR(\Phi_n(t), \rho, \theta)$ is the Signal to Interference Ratio at a given point $(\rho, \theta)$ of the cell in polar coordinates (being an interference limited scenario, noise has been neglected and thus only SIR has been considered). Hence, (5.24) averages the SIR for all points in the network area $A$ covered by the cell.

Considering that any interfering cell $j$ is located, in polar coordinates, at a point ( $d_j$ , $\phi_j$ ) with respect to the reference cell (see Figure 5.8), $SIR(\Phi_n(t), \rho, \theta)$ is written as

$$SIR(\Phi_n, \rho, \theta) = \frac{P_n K_{PL} \rho^{-\chi}}{\sum_{j \in \Phi_n} P_n K_{PL} \rho_j^{-\chi}} . \tag{5.25}$$

for $\rho_{\min} \le \rho \le \rho_{\max}$ and $0 \le \theta < 2\pi$ . $\rho_{\min} > 0$ is a minimum distance between users and the base station due to base station antenna height and $\rho_{\max}$ is the maximum distance to the base station in the cell's coverage area. Constant chunk power $P_n$ is assumed for all chunks and pathloss is modeled as $K_{PL}\rho^{-\chi}$, being $K_{PL}$ a pathloss constant and $\chi$ the pathloss exponent. Since average results in the medium term are of interest, slow and fast varying fading has not been considered in (5.25). The distance between the interfering cell and the point of interest in the reference cell, can be written as $\rho_j = \sqrt{d_j^2 + \rho^2 - 2\rho d_j \cos(\theta - \phi_j)}$ . Then, (5.24) can be simplified to

$$SIR(\Phi_n, \rho, \theta) = \frac{1}{\sum_{j \in \Phi_n} \left(1 + \left(d_j/\rho\right)^2 - 2\left(d_j/\rho\right)\cos\left(\theta - \phi_j\right)\right)^{-\frac{\chi}{2}}} . \tag{5.26}$$

Finally, $\hat{\gamma}_k^{(n)}(t)$ is averaged for all assigned chunks to the cell in a given RL step to obtain $\hat{\gamma}_k(t)$ :

$$\hat{\gamma}_k(t) = \frac{1}{N_k(t)} \sum_{n / y_{kn}(t)=1} \hat{\gamma}_k^{(n)}(t) , \tag{5.27}$$

where $N_k(t) = \sum_{n=1}^{N} y_{kn}(t)$ is the number of assigned chunks for a given action $\Psi_k(t) = (y_{k1}(t), ..., y_{kN}(t))$ of RL-DSA for the $k$ -th cell.

On the other hand, the average user throughput in the cell can be obtained as

$$th_k(t) = \frac{B N_k(t) \hat{\eta}_k(t)}{U_k} , \tag{5.28}$$

where $B$ denotes the chunk bandwidth, $U_k$ is the average number of users in the cell, and $\hat{\eta}_k(t)$ is an estimation of the average cell spectral efficiency for a given spectrum assignment. Similarly to $\hat{\gamma}_k(t)$ ,

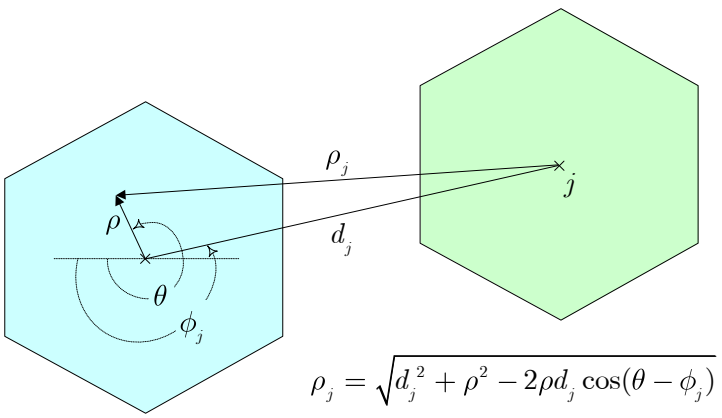


**Figure 5.8 Illustrative layout for Signal-to-Interference (SIR) computation.**

$\hat{\eta}_k(t)$ can be obtained as

$$\hat{\eta}_k(t) = \frac{1}{N_k(t)} \sum_{n/y_{kn}=1} G(U_k, \hat{\gamma}_k^{(n)}(t)) \iint_A \frac{1}{A} q\left(SIR(\Phi_n(t), \rho, \theta)\right) \rho d\rho d\theta \tag{5.29}$$

where $q\left(SIR(\Phi_n(t), \rho, \theta)\right)$ is the spectral efficiency in bits/s/Hz for a given value of the SIR. For instance, function $q$ can be the mapping table given in Table 4.1 in Chapter 4. $G(U, \hat{\gamma}_k^{(n)}(t))$ is a gain factor that captures the multiuser diversity features of the short-term scheduling strategy used in the cell as a function of the average number of users $U_k$ and the average estimated SINR per chunk $\hat{\gamma}_k^{(n)}(t)$. In particular, it is well known that the achieved spectral efficiency for a Round Robin short-term scheduling strategy does not depend on the number of users in the cell, because RR leads to equal users' transmission probability. Then, for a RR strategy a proper setting is $G(U_k, \hat{\gamma}_k^{(n)}(t)) = 1$ for all $U_k$ and $\hat{\gamma}_k^{(n)}(t)$.

However, a channel-aware scheduler with unequal users' transmission probabilities, such as Proportional Fair, leads to a dependence of the achieved spectral efficiency with the number of users in the cell under certain SINR patterns. Hence, for PF, gain factor concept was developed by recent studies [74][75]. This gain factor $G(U, \hat{\gamma}_k^{(n)}(t))$ depends on the number of users and the SINR distribution over the cell, as shown in Figure 5.9. This figure plots a set of PF gain factor curves obtained for exhaustive simulations that focus on the central cell in a two ring macrocell scenario with different intercell interference patterns leading to different average SINRs per chunk $\hat{\gamma}_k^{(n)}$.

This gain factor depends on the number of users and average cell SINR. Notice that no gain (gain value equal to 1) is obtained for a single user since it is not possible to exploit multiuser diversity in this case. Figure 5.10 compares cell spectral efficiency results obtained through a simulation and the estimation model given in (5.29), for RR and PF schedulers respectively, showing in each case a very satisfactory matching.

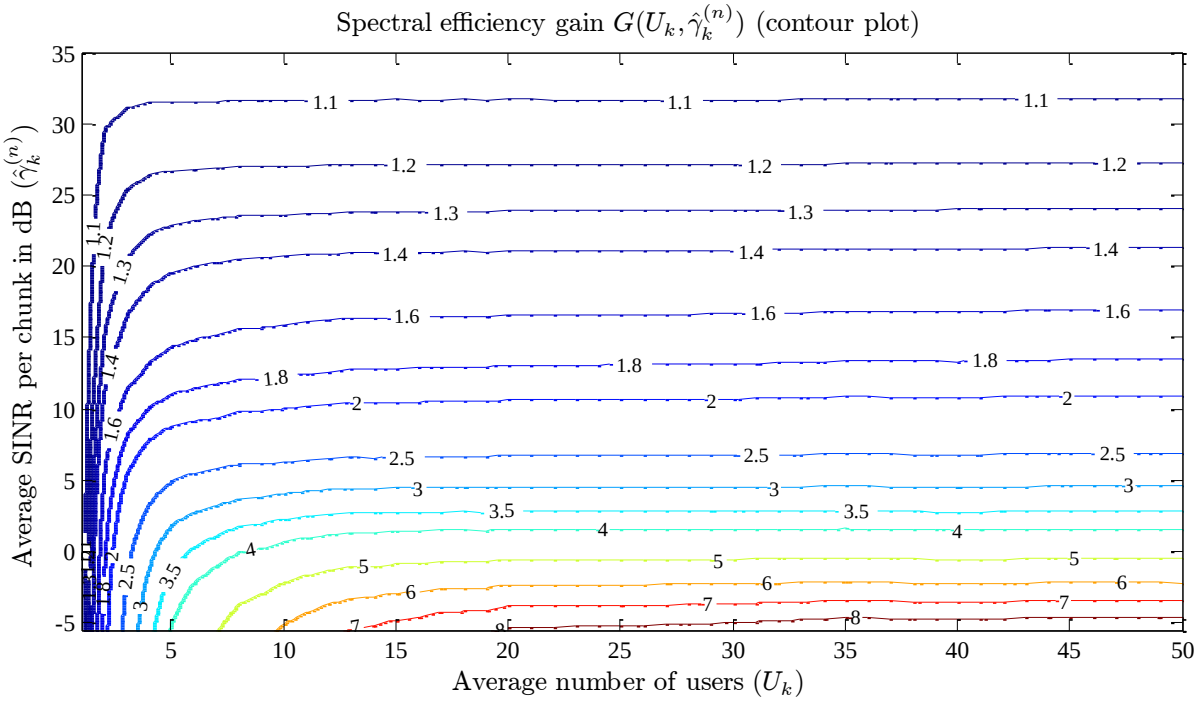


**Figure 5.9 Spectral efficiency gain versus average SINR and number of users (contour plot).**

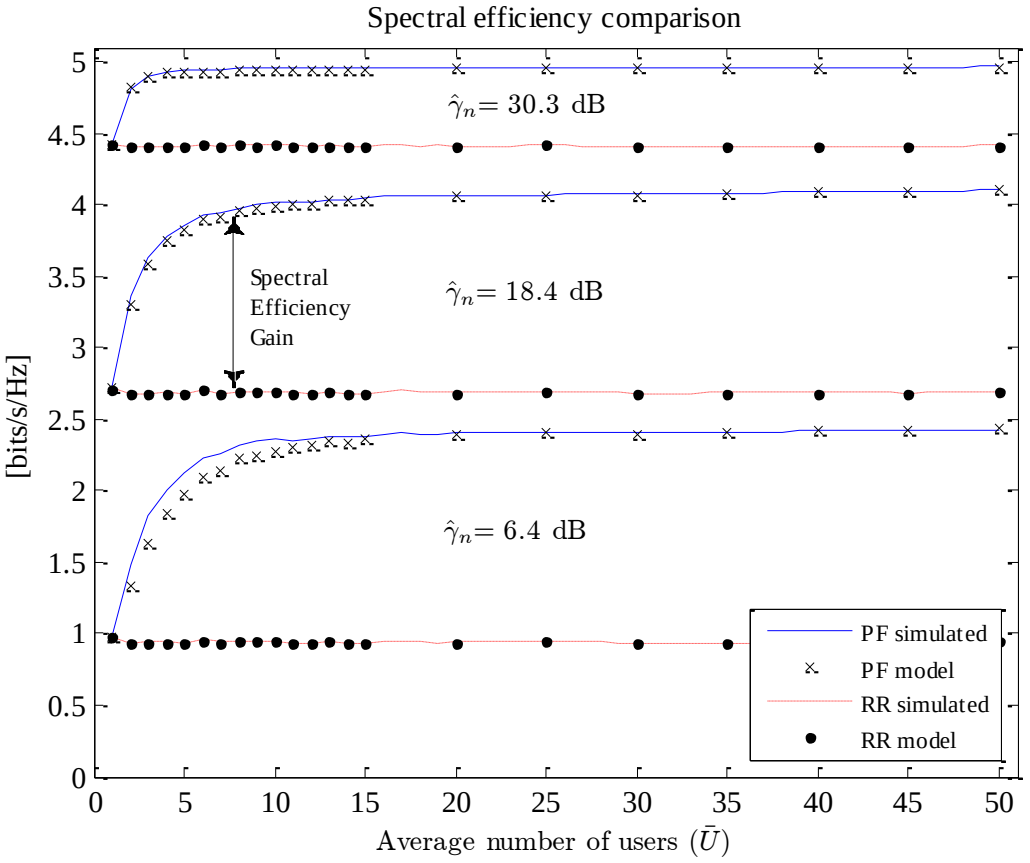


**Figure 5.10 Spectral efficiency comparison for RR and PF schemes and selected average SINR values**

### 5.3.2.2 Non-cooperative Decentralized DSA framework

In the case of non-cooperative spectrum assignment scheme the average SINR and user throughput are estimated from measurements, because the spectrum usage of adjacent cells is unknown. Let $s_{\bar{\gamma}_k^{(n)}}(\bar{\gamma}_k^{(n)})$ be the average SINR PDF for chunk $n$ in cell $k$ computed by Status Observer from users' measurement reports. The average cell SINR per chunk $\hat{\gamma}_k^{(n)}(t)$ can be obtained as in (5.23). Then, the average cell SINR is computed as

$$\hat{\gamma}_k(t) = \frac{1}{N_k(t)} \sum_{n/y_{kn}(t)=1} \hat{\gamma}_k^{(n)}(t). \tag{5.30}$$

On the other hand, an estimation of the average spectral efficiency can be obtained as

$$\hat{\eta}_k(t) = \frac{1}{N_k(t)} \sum_{n/y_{kn}(t)=1} G(U_k, \hat{\gamma}_k^{(n)}(t)) \int_{-\infty}^{\infty} q(\bar{\gamma}_k^{(n)}) s_{\bar{\gamma}_k^{(n)}}(\bar{\gamma}_k^{(n)}) d\bar{\gamma}_k^{(n)}, \tag{5.31}$$

where, as in detailed in subsection 5.3.2.1, $G(U_k, \hat{\gamma}_k^{(n)}(t))$ is the spectral efficiency gain factor that depends on the underlying short-term scheduling strategy. Finally, the average user throughput in the cell can be estimated by following (5.28), and considering (5.31).

## 5.4 Summary

This chapter has presented the DSA algorithms proposed in this thesis that run in the DSA framework introduced in previous Chapter 4. The aim of the algorithms is to find the best OFDMA frequency chunks that should be assigned to each cell in order to improve the spectral efficiency maintaining certain QoS requirements and adapt to heterogeneous distributions of the traffic load, surpassing then the limitations of frequency reuse factors.

The HEUR-DSA algorithm is a simple but effective algorithm that assigns chunks to cells according to intuitive assignment rules that determine (*i*) the number of chunks that should be granted to a cell depending on its traffic load and (*ii*) the specific set of chunks that mitigate the intercell interference and thus enables an improvement of the spectral efficiency (i.e., the throughput per unit of bandwidth). Four flavors of the algorithm have been proposed with the abovementioned procedure showing different approaches to perform dynamic spectrum assignment. Then, for instance, distinction between inner and outer users can be done to improve interference conditions of edge (outer) users with HEUR-DSA3 and HEUR-DSA4 while still maintaining certain adaptability of the assignment scheme. Preliminary validation results showed that HEUR-DSA algorithm certainly vary the number of chunks per cell in accordance to the traffic load per cell.

On the other hand, a machine learning strategy based on reinforcement learning has been proposed with the RL-DSA algorithm. Reinforcement learning methods are characterized by low complexity and optimal behavior in the sense that a maximum reward is guaranteed in the long-term. Thus, differently from HEUR-DSA, RL-DSA assures that the optimal spectrum assignment can be found for a certain number of samples of the reward signal. Concretely, each action of RL-DSA represents a candidate spectrum assignment. Then a reward is obtained for each candidate spectrum assignment, and RL converges to the spectrum assignment that maximizes reward. Then, the key for radio resource performance optimization is to define the reward in terms of certain performance metrics. Particularly, two reward signals per cell have been proposed. One devoted to spectral efficiency optimization and the other to SINR optimization whereas in both cases the QoS satisfaction of the users is guaranteed. Validation results demonstrate an appropriate convergence of the algorithm in a simple but controlled scenario and an interesting low complexity and high robustness in the convergence to the optimal solution despite the initial configuration of the RL-DSA parameters.

Finally, this chapter has finished with a description of the implementation of the Network/Cell Characterization entities in proposed framework. This entity supports the execution of the DSA algorithms. In the case of HEUR-DSA the Network/Cell Characterization entity returns the cost of assignment of a given chunk to a cell, which HEUR-DSA uses to decide the best (lowest cost) chunks to assign. On the other hand, in the case of RL-DSA the Network/Cell Characterization entity builds the reward signal for each candidate spectrum assignment of RL-DSA. The proposed models are simple to implement and then quickly respond to DSA algorithms requests.

# 6

# EVALUATION METHODOLOGY

## Outline

# 6 Evaluation methodology

This chapter is devoted to present the evaluation methodology followed in this work to obtain the performance comparison results later discussed on Chapter 7. The chapter is organized in three parts. In the first part, the simulation tool developed and employed to execute dynamic simulations is described. This tool is used to simulate a downlink OFDMA cellular scenario where the performance for different spectrum assignment strategies included the ones proposed in previous Chapters 4 and 5 is obtained. The second part of this chapter is devoted to define the performance metrics used to compare those spectrum assignment strategies. Finally, the third part of this chapter describes application scenarios where the performance comparison results of Chapter 7 have been obtained.

## 6.1 Simulation tool

A simulation tool has been developed to evaluate different spectrum assignment strategies in a multicell scenario based on a downlink Orthogonal Frequency Division Multiplexing (OFDMA) radio access technology. It is a simulator where both dynamic (i.e. including mobility, traffic generation, etc.) and quasi-static (i.e. based on snapshots, without user mobility) simulations can be performed.

Macrocells and femtocells can be available in the scenario. Macrocells are deployed with a canonical hexagonal layout and omnidirectional antennas. Femtocells also use omnidirectional antennas but can be deployed inside buildings or at randomly positions in a certain area.

The simulator can compare fixed spectrum assignment strategies with the proposed DSA framework and algorithms in Chapter 4 and Chapter 5 respectively, in terms of several performance metrics such as, for instance, spectral efficiency, user's QoS, or spatial spectrum usage over the scenario.

The simulation tool is a stand-alone command window application with a modular design implemented in C/C++ programming language. In the following, inputs, functional architecture, simulation procedure, implemented models, and outputs for the simulation tool are specified.

### 6.1.1 Inputs

The following Table 6.1 collects main inputs that are available for the simulator regarding different aspects such as network deployment, traffic characterization, propagation, mobility, packet scheduling, and spectrum management.

These inputs are provided in configuration text files. Figure 6.1 shows the format of the configuration files, where each line represents the values for a given input variable.

```
//GLOBAL VARS
//Format [var_name]=[var_value][Tab][Any comment up to 99 char]
//SIMULATOR
results_dir=.\Results\
config_dir=.\Config\
```

**Figure 6.1 Format of the configuration files**

**Table 6.1 Inputs for DSA-OFDMA simulator**

| Type of input | Input | Comment |
|---|---|---|
| Specific Simulator Configuration parameters | Frame time | *In seconds* |
| | Simulation time | *In frames* |
| | Seed | *Random seed value* |
| Cellular deployment | Number of base stations | *[1..19]* |
| | Position of base stations | *In a text file* |
| | Cell radii | *In km* |
| Radio parameters | Maximum transmit power per cell | *In dBm* |
| | Noise floor power density | *In dBm/Hz* |
| | Number of OFDMA chunks in the system | *[1-48]* |
| | Chunk bandwidth | *In Hz* |
| | Carrier frequency | *In Hz* |
| | Modulation and coding schemes | *See 6.1.4* |
| Traffic | Number of users | *Minimum one user per cell* |
| | Traffic models | *Full-buffer (see 6.1.4)* |
| | Spatial distribution of the users | *Input file with the percentage of the traffic load per cell is given. Also, a given distribution can be changed during simulation.* |
| Propagation | Path loss | *Exponent, minimum coupling loss and wall penetration losses* |
| | Shadowing | *Standard deviation and decorrelation distance* |
| | Fast Fading | *Power delay profile* |
| Mobility | Users' speed | *In km/h* |
| | Trajectories | *Optionally input file with predefined trajectories (test purposes)* |
| Short term scheduling | Strategy | *Round Robin, Proportional fair* |
| | Specific parameters of the strategy | *Averaging window* |
| Spectrum management | Strategy | *Fixed (FRF1, FRF3, PR, SR, …)*<br>*HEUR-DSA*<br>*RL-DSA* |
| | Specific parameters of the strategy | *e.g. type of framework reward parameters, …* |

## 6.1.2 Functional architecture

The DSA-OFDMA simulator has the functional architecture depicted in Figure 6.2. Inputs given to configure scenario deployment, users' traffic generation, users' mobility, propagation models, scheduling, and spectrum are used to characterize the users' trajectories, users' path-gains, traffic buffers status, and to execute the short-term scheduling and DSA algorithms.

The simulator is compound of different functional blocks described in the following.

- Trajectories Functional Block:

  This block is in charge of computing the position of the users in the scenario. The coordinates of the users with respect to all cells is calculated. Also, if dynamic simulations are configured, users are moved according to their speed and the mobility model.

- Path Loss Computations Functional Block:

  Taking into account the relative position of users respect to cells, the distance dependant path loss can be computed. Also in this block, the slow fading and fast frequency dependant fading is calculated according to propagation models defined in subsection 6.1.4.

- Link Level Characterization Functional Block:

  This block computes the Signal-To-Interference plus Noise (SINR) ratio for each user in each one of the available chunks in its serving cell. After that, it is possible to estimate the Modulation and Coding Scheme (MCS) that each user could use in each chunk from link-level curves. In other words, it is possible to estimate the bit rate that a user could transmit if short-term scheduler grants him/her the channel (i.e. a specific chunk in a frame).

- Traffic Generation Functional Block:

  Depending on the traffic model considered for the users, data packets are generated and stored in buffers. One independent buffer per user is supposed in the base station.

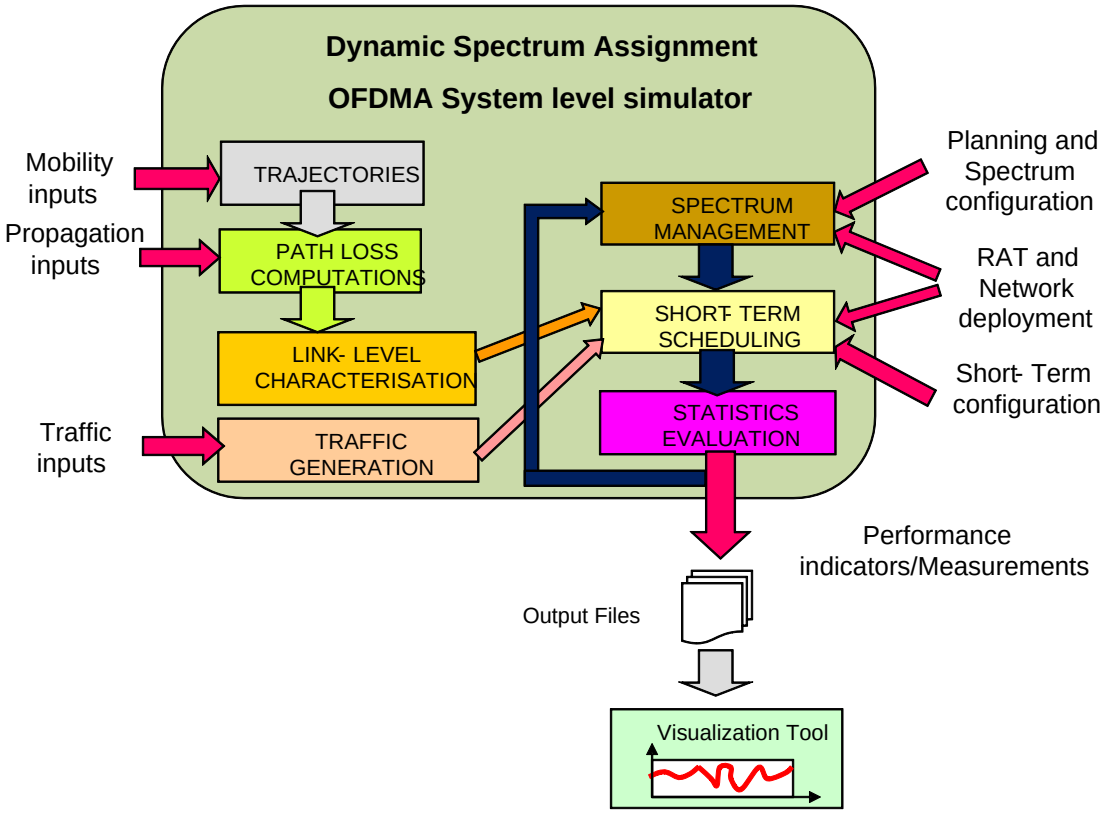


**Figure 6.2 Functional Architecture of DSA-OFDMA simulator**

- Spectrum Management Functional Block:

  This block is responsible of executing the different dynamic spectrum assignment strategies envisaged in the simulator. As inputs it considers configuration values such as the available system bandwidth or specific DSA strategies parameters. It also uses on-line statistics from the simulator such as the current load per cell, or QoS indicators. The output of this block is a cell-by-cell spectrum assignment that determines the available chunks per cell to schedule in the short-term user's transmissions.

- Short-Term Scheduling Functional Block:

  Here the short-term scheduling strategy is executed and users' transmissions are scheduled frame-by-frame into the available chunks in each cell. Also this block updates the users' buffers accordingly to the resources scheduled per user.

- Statistics Evaluation Functional Block:

  This block performs the on-line computation of several statistics. Some of them are for internal use of the simulator and others are stored at the end of the simulation into plain text files that can be processed and visualized off-line. The visualization tool is not part of the simulator.

### 6.1.3 Simulation procedure

The DSA-OFDMA simulator executes a loop that continuously updates several aspects of the simulator as it is shown in Figure 6.3. The following procedures can be defined:

- Simulator initialization:

  This procedure executes once at the beginning of the simulation. It loads all the configuration parameters and based on them builds the scenario. For instance, for the network deployment, base-station C++ objects are created and each one is configured to be in a specific position in the cellular grid. Also, initial spectrum assignment is set accordingly to the selected spectrum assignment strategy selected.

After initialization, the simulation loop is started and executed until the simulation time is reached where each iteration or loop represents a frame. Within the loop, three main procedures are executed: Update Scenario, Update Statistics and DSA.

- Update Scenario:

  This procedure updates users' propagation conditions, estimates the link-level status and thus computes the users' available rate in a frame and re-generates the users' traffic depending on their traffic models. After that, short-term scheduling is performed in each cell, what determines the users that transmit in the frame. After that, the users' buffers are updated accordingly to the granted rate by the short-term packet scheduling and link adaptation.

- Update Statistics:

  This procedure updates the statistics relative to the whole system, each cell, and each user, taking into account current status in the frame.

- DSA:

  The simulator checks continuously if a new spectrum assignment is needed taking into account the system and cell status statistics. If so the DSA algorithm (HEUR-DSA or RL-DSA) is executed.

Finally, once the simulation loop has completed, the simulator saves the statistics into files to be analyzed afterwards and closes.

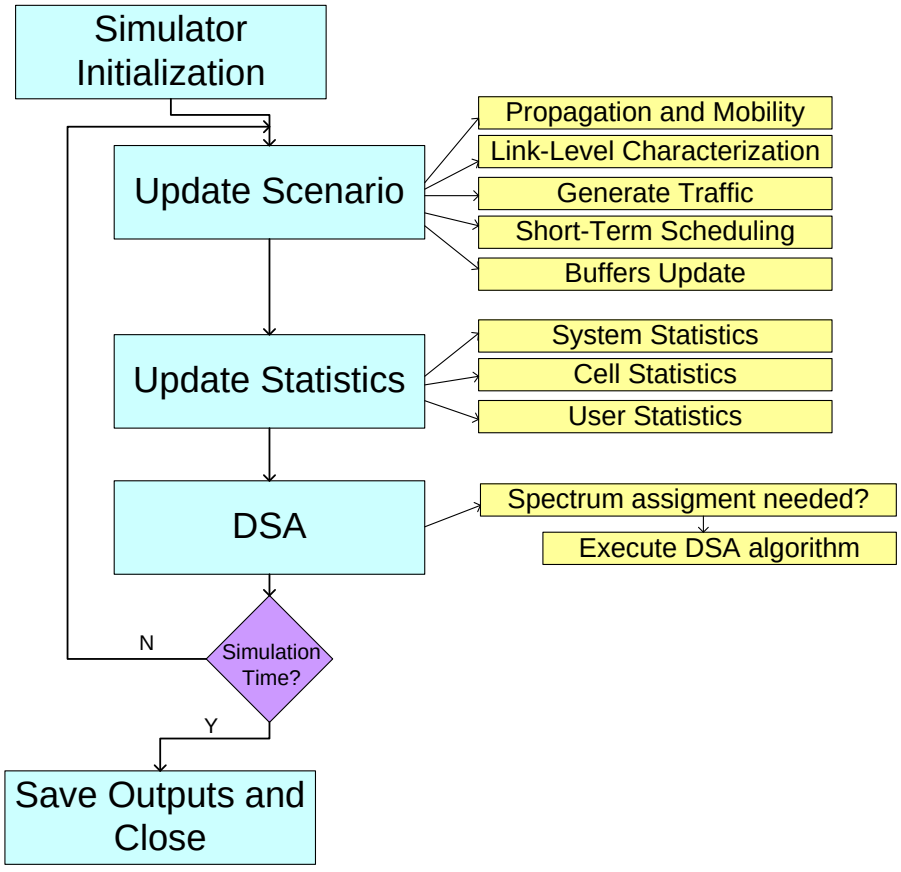


**Figure 6.3 DSA-OFDMA simulator simulation loop**

## 6.1.4 Models and algorithms

### 6.1.4.1 Propagation Models

Distance dependant path loss, slow fading and fast fading are considered for all links between a given base station and all users in the scenario. In a given instant, path loss and slow fading are considered the same for all frequency chunks whereas fast fading varies between chunks.

- Distance Dependant Path Loss model

  Path loss is computed for all users respect all base stations. The path loss expression is [76]:

$$PL = \max\left(MCL, 128.1 + 37.6\log_{10}(d)\right) \text{ dB}, \tag{6.1}$$

  where $d$ is the user-base station distance in km and MCL represents the minimum coupling loss that is set to 70 dB. This expression is valid for urban and suburban environments with a carrier frequency of 2GHz.

  The aforementioned path loss model is employed for outdoor links between macrocells and users. For femtocells, indoor propagation is modeled with COST 231multi-wall model [77] without the effect of floor attenuation (since only buildings with one floor are considered). Then,

$$PL = \max(MCL, 37 + 32\log_{10}(d/1000) + L_{wall}n_{wall}) \text{ dB}, \tag{6.2}$$

  where $d$ is the distance in km, $n_{wall}$ are the number of walls that signal has to cross and $L_{wall}$ is the penetration loss of each wall. Inner walls are considered as 'narrow' walls (i.e., with small penetration losses), whereas external walls are considered as 'thick' walls. Penetration losses for doors and windows are also considered. A MCL of 30 dB is considered in this case.

  The COST 231multi-wall model is only used for links between a femtocell access points and users belonging to that femtocell. In any other case (i.e., links between macrocells and femtocells users, between femtocells and macrocell users and between femtocell and other femtocell users) the macrocell PL plus appropriate penetration losses is considered.

- Slow Fading (Shadowing) model

  The long-term fading is modeled as a Gaussian distribution with standard deviation $\sigma_s$ in dB around the mean path loss in dB (lognormal fading). Gudmundson correlation model is employed [76].

- Fast Fading model

  ITU pedestrian A tapped-delay-line model is employed to simulate fast frequency dependant fading [76]. Fast fading is updated in each iteration of the simulation loop, that is, in each frame.

#### 6.1.4.2 Link-Level Characterization

Users' transmission bit rate is variable by means of Adaptive Modulation and Coding (AMC). As given in Chapter 4, an analytic and a table-based approach have been considered to model AMC, which are reproduced here.

The analytic approach is based on the following formula [68] for a target Bit Error Rate (BER) and Rayleigh fading channel:

$$q_{m,n}(f) = \log_2\left(1 + \frac{-1.5\gamma_{m,n}(f)}{\ln(5BER)}\right), \tag{6.3}$$

where $q_{m,n}(f)$ is the spectral efficiency in bits/s/Hz attained in the frame $f$, BER<(1/5)exp(-1.5) ≈ 4.46% and $\gamma_{m,n}(f)$ < 30 dB.

On the other hand, table-based AMC is based on the following table in the ambit of 3GPP LTE systems [69]:

**Table 6.2 Adaptive modulation and coding**

| SINR threshold [dB] | Modulation *m* [bits/s/Hz] | Coding Rate *r* | Achievable spectral efficiency *q=m*r* [bits/s/Hz] |
|---|---|---|---|
| < 0.9 | - | - | 0 |
| ≥ 0.9 | 2 (QPSK) | 1/3 | 0.66 |
| ≥ 2.1 | 2 (QPSK) | 1/2 | 1 |
| ≥ 3.8 | 2 (QPSK) | 2/3 | 1.33 |
| ≥ 7.7 | 4 (16QAM) | 1/2 | 2 |
| ≥ 9.8 | 4 (16QAM) | 2/3 | 2.66 |
| ≥ 12.6 | 4 (16QAM) | 5/6 | 3.33 |
| ≥ 15.0 | 6 (64QAM) | 2/3 | 4 |
| ≥ 18.2 | 6 (64QAM) | 5/6 | 5 |

#### 6.1.4.3 Mobility Models

- Random Walk as in [76] for pedestrian and vehicular models in urban area. In this model users' trajectories randomly change direction at each update with a certain probability (20%). The maximum variation in the direction is limited to 45 degrees.

- Predefined Trajectories: For testing purposes there is also the possibility of configure in a file the trajectories of some of the users.

#### 6.1.4.4 Traffic Models

Full-buffer traffic model is implemented as specified in [78]. This traffic model assumes that users always have packets to transmit in their queues. This traffic model has been selected to perform comparison between tested spectrum assignment strategies in terms of throughput and spectral efficiency. It is worth to mention that delay constraints have not been considered.

#### 6.1.4.5 Short-Term Scheduling Strategies

Short-term packet scheduling strategies presented in Chapter 4 have been implemented. These are:

- Round-Robin: Round Robin assigns the channel cyclically to users, but its performance is poor since it is not channel aware in its decisions.

- Proportional-Fair: Proportional Fair (PF) [70] has been proposed to provide a trade-off between fairness and system throughput and to exploit multiuser diversity in time.

#### 6.1.4.6 Spectrum Assignment Strategies

Several spectrum management strategies can be selected in the simulator to be evaluated. They include fixed strategies that adopt a spectrum assignment at the beginning of the simulation that remains unaltered until the end of the simulation. They deploy different frequency reuse factors (FRF) in the scenario [40]. Possible fixed strategies are FRF1, FRF3, Partial frequency Reuse [42] and Soft frequency Reuse [44]. On the other hand, dynamic strategies that vary the cell-by-cell spectrum assignment during simulation have been implemented. Heuristic Algorithms and Reinforcement Learning Algorithms, as detailed in Chapter 5 are available.

### 6.1.5 Outputs

Several statistics are monitored on-line during simulator execution. These statistics are saved into several trace files at the end of the simulation for post-processing and visualization of the results. For instance, these trace files can be loaded and processed using Matlab®.

For any monitored parameter the following statistics could be available:

- Minimum Value
- Maximum value
- Average Value
- Variance
- Probability Density Function (p.d.f.)
- Cumulative Probability Function (c.d.f.)
- 5 and 95 Percentiles

Any other statistic can be easily incorporated by means of a home-made C++ class that provides the necessary methods to compute the statistic values. Figure 6.4 shows sample of an output trace file.

```
//min value     max value  mean  variance   5thpercentile   95thpercentile  others
0.000000   291916.167665   221029.367212   37284.207552    150000.000000   170000.000000   0.026667
0.000000   283687.500000   219440.845868   41805.449019    135000.000000   165000.000000   0.043333
0.000000   288750.000000   223879.291398   43598.102963    135000.000000   150000.000000   0.030000
0.000000   290000.000000   218515.130651   41897.692930    135000.000000   155000.000000   0.026667
0.000000   280500.000000   218032.034584   41872.262364    150000.000000   160000.000000   0.013333
0.000000   282400.000000   223567.687350   34844.309980    170000.000000   180000.000000   0.000000
0.000000   284700.000000   225850.095488   40512.759839    135000.000000   155000.000000   0.033333
0.000000   289166.666667   216767.420510   49817.509088    135000.000000   140000.000000   0.030000
0.000000   282075.000000   224731.098425   36121.861584    150000.000000   170000.000000   0.000000
0.000000   285000.000000   216258.904779   51999.707036    120000.000000   140000.000000   0.070000
0.000000   278750.000000   229056.537907   32826.281633    165000.000000   180000.000000   0.000000
0.000000   277875.000000   228340.053969   36759.670989    170000.000000   180000.000000   0.000000
0.000000   288214.285714   219872.413819   40618.816969    145000.000000   150000.000000   0.006667
0.000000   285000.000000   223819.225360   44252.015453    155000.000000   165000.000000   0.003333
```

**Figure 6.4 Sample of a statistics trace file**

## 6.2 Key Performance Indicators

Main performance metrics or Key Performance Indicators (KPIs) used for results of the thesis in Chapter 7 are defined here. Nevertheless, some well known metrics such as cumulative distribution functions of the SINR, spatial SINR patterns over the scenario, or cell/user average throughput can be also used in some results. All the KPIs are averaged during a time window of $l$ frames.

### 6.2.1 Dissatisfaction probability

This metric reflects the probability that the average user throughput is below a given target. Hence, it is used as an indicator of users' QoS performance. The average dissatisfaction probability per cell is formally defined as

$$P_k^{th_{\text{target}}} = \frac{1}{l}\sum_{f=1}^{l}\frac{1}{U_k}\sum_{m=1}^{U_k}\theta_m^{th_{\text{target}}}(f). \tag{6.4}$$

$th_{\text{target}}$ is the user satisfaction throughput, $U_k$ is the number of users in cell $k$ and $\theta_m^{th_{\text{target}}}(f)$ is an indicator per the $m$-th user and the $f$-th frame defined as

$$\theta_m^{th_{\text{target}}}(f) = \begin{cases} 1 & th_m(f) < th_{\text{target}} \\ 0 & th_m(f) \geq th_{\text{target}} \end{cases}, \tag{6.5}$$

where $th_m(t)$ is the average throughput that user $m$ experiments during last second before frame $f$.

### 6.2.2 Throughput Fairness

It reflects the balance between the throughput performance of users located close to the base station and users at the edge of the cell. Here the definition of throughput fairness given in [79] is taken. It is defined in terms of the Normalized Throughput Bound (NTB), which is defined as a straight line of slope 1 in the plot of the cumulative distribution function (CDF) of the normalized user throughput (normalized to average user throughput). Fairness for a given spectrum assignment strategy increases as its obtained normalized user throughput CDF lies to the right of the NTB. Notice that point $(x, x)$ in the NTB implies that at least $(1-x)\cdot 100\,\%$ of users have at least $x\cdot 100\,\%$ of the average user throughput. Figure 6.5 illustrates this concept, where the blue line stands for a possible normalized throughput CDF.

In some cases the 5th percentile of the average user throughput per cell normalized to mean throughput is taken as a fairness criterion. This value is also at the right of the normalized throughput bound.

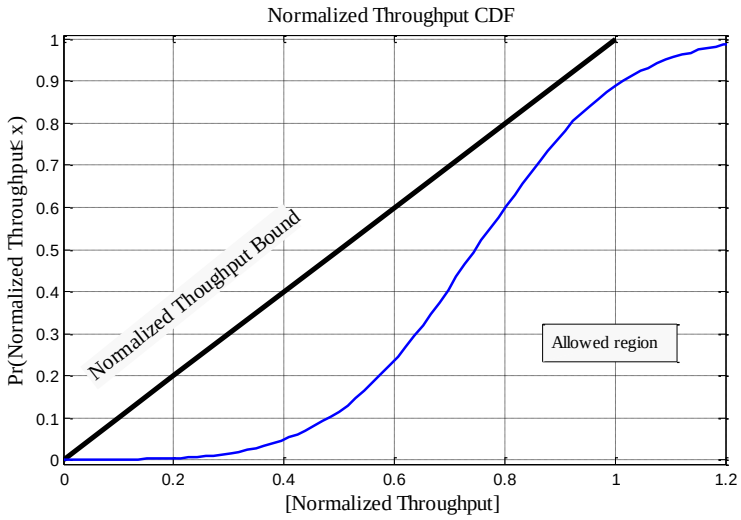


**Figure 6.5 Normalized Throughput Bound concept illustration**

### 6.2.3 Spectral efficiency

Spectral efficiency defines the normalized number of bits per unit of spectrum in a given cell $k$. It is defined as

$$\eta_k = \frac{1}{l}\sum_{f=1}^{l}\frac{TH_k(f)}{W_k(f)}, \tag{6.6}$$

where $TH_k(f)$ is the aggregate throughput of all users in cell $k$, and $W_k$ is the bandwidth given to that cell in Hz. Spectral efficiency is measured in bits/s/Hz.

### 6.2.4 Useful Released Surface

Finally, to measure the capability of releasing spectrum in a given band for secondary cognitive radio usage the new metric called Useful Released Surface (URS) presented in [80] is retained. The URS defines the surface where a given bandwidth can be used by secondary cognitive radio users respecting primary users' maximum interference level constrains. Formally, URS is defined as

$$URS = \sum_{n=1}^{N} B^{(n)} \sum_{a=1}^{\Lambda_n} S_a^{(n)} \omega_a^{(n)} \quad \text{MHz} \times \text{km}^2, \tag{6.7}$$

where $B^{(n)}$ is the bandwidth of the $n$-th chunk, and $\Lambda_n$ is the set of non-contiguous areas where chunk $n$ can be released. $S_a^{(n)}$ is the surface of each one of the areas, which is computed as the surface of all contiguous cells that do not use chunk $n$ minus the surface of a protection area around cells that do use the $n$-th chunk to protect from interference primary users in them. Notice that with this definition, only the areas where secondary users are allowed to use the spectrum without harming primary users' transmissions are considered. Here the first tier of neighboring cells around a cell has been taken as a protection area. Finally, $\omega_a^{(n)}$ is a weighting factor for each area depending on the number of secondary cognitive radio users that exist on that area. For instance, $\omega_a^{(n)}$ can be set to the fraction of potential secondary users that would exist in an area assuming that they are uniformly distributed in the scenario

## 6.3 Application scenarios

Four different scenarios have been set up to obtain performance comparisons between state-of-the-art spectrum assignment strategies and the DSA framework proposed in this thesis.

### 6.3.1 Scenario A. Static distribution of the traffic load in macrocell deployment

The purpose of this scenario is to compare the behavior of the spectrum assignment strategies under different spatial distributions of the traffic load over a downlink macrocell cellular layout. Hence four spatial distributions of the load over a cellular deployment with 19 omnidirectional cells were simulated as shown in Figure 6.6. Figure 6.6 (a) depicts a homogeneous distribution and Figure 6.6(d) a highly heterogeneous distribution, where Figure 6.6(b) and Figure 6.6(c) depict intermediate situations. For each case, the load percentage per cell and the cell number is included in the figure. Moreover, two Cells-of-Interest (CoI) are highlighted to evaluate the results per cell. CoI#3 represents the most loaded cell in

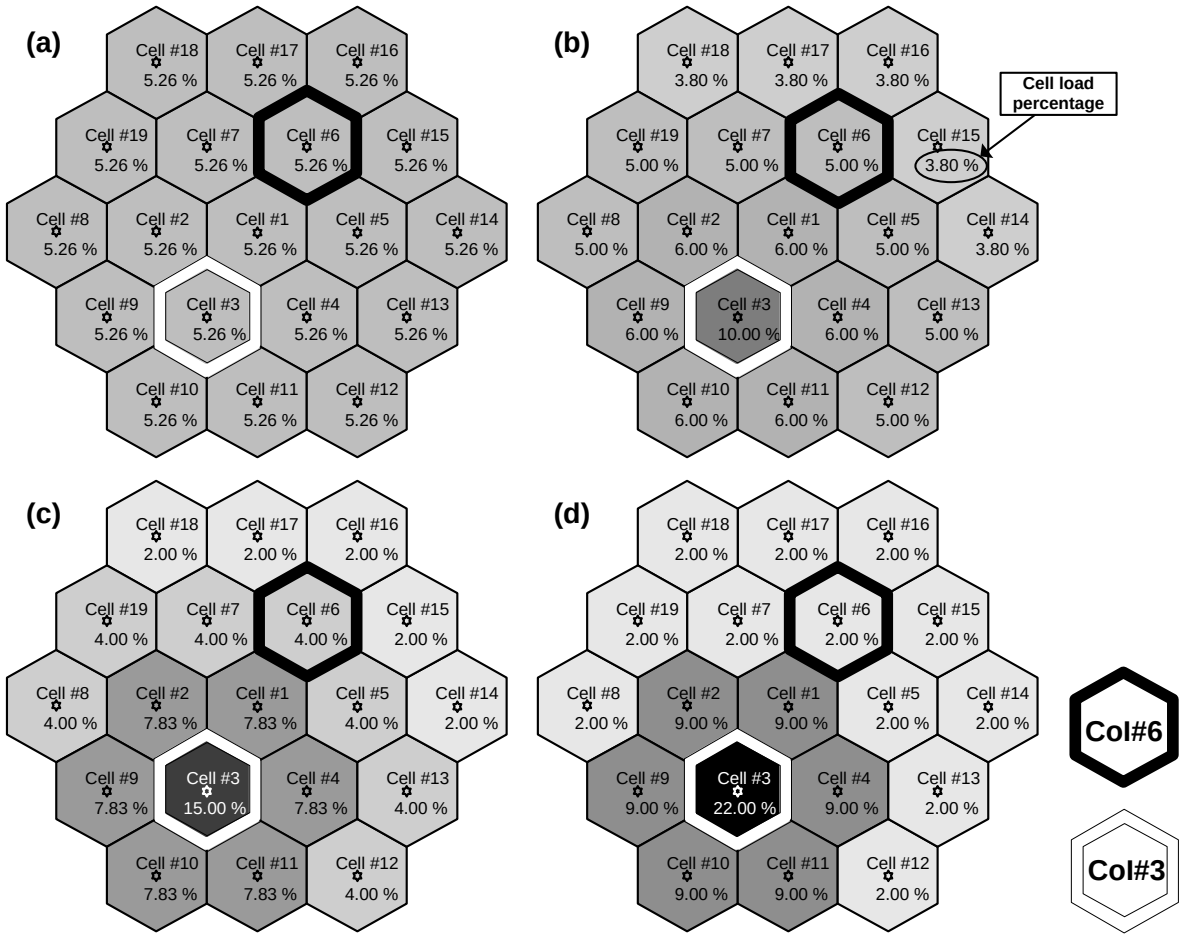


**Figure 6.6 Four spatial distributions of the load simulated in Scenario A**

heterogeneous distributions. On the other hand, CoI#6 represents a cell whose load percentage decreases throughout the patterns as load concentrates in CoI#3.

Users are uniformly distributed within a cell and their mobility is restricted to the cell where they belong to in order to maintain the load ratio between cells, (i.e., no handover effects are modeled). The Short-Term Scheduler employed is a PF scheduler. Only one user is scheduled per chunk in each frame although a user could get more than one chunk per frame. For the link adaptation, the analytical Adaptive Modulation and Coding scheme given in section 6.1.4.2 is employed with a maximum spectral efficiency limited to $\eta_{\max}$ =4 bits/s/Hz and a target BER of $10^{-3}$.

Users are deployed in the scenario with their buffers always full and all users request a satisfaction throughput $th_{\text{target}}$ =128Kbps. This means that a user has always information to transmit and then, he/she aims to get as much capacity as possible above 128Kbps. If available, edge users' transmissions have priority over edge chunks. Then the scenarios are under the worst traffic load given a number of users in the scenario, which is equal to 300 active users in all test performed.

The maximum number of chunks in the system is $N = 12$. Hexagonal cells of radius $R = 0.5$ Km are employed. The maximum power per chunk is $P_{\max}$ =32.21 dBm (i.e., the maximum power per cell is $N \times P_{\max} = 43$dBm) what assures proper coverage levels at the maximum cell radius.

Finally, Table 6.3 collects the simulation parameters for Scenario A.

**Table 6.3 Simulation parameters for Scenario A**

| | |
|---|---|
| Frame duration | 2 ms |
| Simulated time | 1 hour |
| KPIs averaging time | $l$ =2500 frames |
| Number of cells | $K$ = 19 |
| Cell Radius | $R_c$ = 500 m |
| Number of antennas | 1 Tx, 1 Rx |
| Antenna Patterns | Omnidirectional |
| Maximum Power per chunk | $P_{max}$ = 32.21 dBm |
| Carrier Frequency | 2 GHz |
| Number of Chunks | $N$ =12 |
| Chunk bandwidth | 375 kHz |
| Path Loss at d km [dB] | $128.1+37.6\log_{10}(d)$ |
| Path Loss Exponent | $\chi$ = 3.76 |
| Shadowing | Lognormal |
| Standard deviation | $\sigma_s$ =8 dB |
| De-correlation model | [76] |
| De-correlation distance | 5 m |
| Small Scale Fading Model | ITU Ped. A [76] |
| UE thermal noise | -174 dBm/Hz |
| UE noise factor | 9 dB |
| UE speed | 3 km/h |
| UE Satisfaction throughput | $th_{target}$ =128 kbps |
| Number of users | 300 users |
| Scheduling | Proportional Fair |
| Averaging Window size | $T_W$ = 50 frames |
| Adaptive Modulation and Coding | Analytic (see section 6.1.4.2) |
| Target BER | $10^{-3}$ |

### 6.3.2 Scenario B. Dynamic distribution of the traffic load in macrocell deployment

This scenario aims at evaluating the performance of considered spectrum assignment strategies under a dynamic distribution of the traffic load in a macrocell deployment. Hence, how the different strategies adapt to traffic variations is examined. To this end, a similar deployment to that in Scenario A is employed, where the macrocell layout is composed of 19 hexagonal cells.

Then, the performance of the system is evaluated during one hour to capture changes in the spatial distribution of the load (users). Particularly, a temporal evolution of the traffic load per cell is enforced. In that respect, three types of cells can be distinguished in the scenario, as it is shown in Figure 6.7. At the beginning all cells are equally loaded with 15 users (i.e., 285 users in total). After 25 minutes, each type of cell varies its number of users in $u_{type1}$, $u_{type2}$ and $u_{type3}$ users per minute respectively. These variations take place only during a 10 minutes period between minutes 25 to 35. Therefore, after that period of time users are heterogeneously distributed among cells. Two different variations have been considered to show the performance of different distributions. On the one hand, $u_{type1}$=2, $u_{type2}$=1 and $u_{type3}$=-1 leads to a distribution where approximately 15% of the users are in the central cell. On the other hand, $u_{type1}$=6, $u_{type2}$=0 and $u_{type3}$=-1 leads to a distribution where 1/3 of the users are in the central cell.

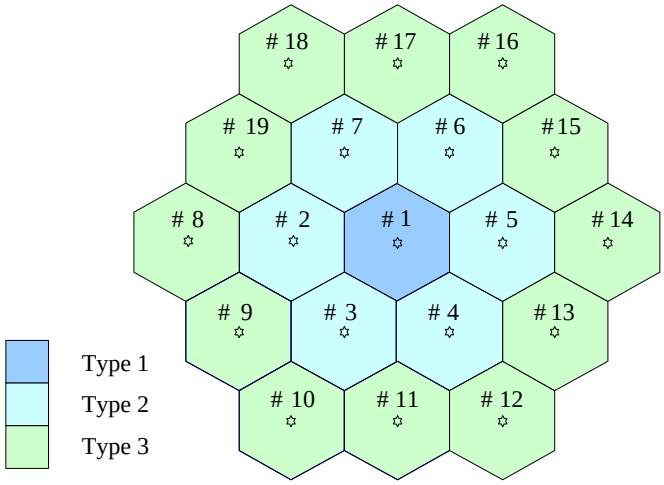


**Figure 6.7 Scenario B layout for simulation.**

Users are distributed homogeneously within a cell, and they move at the speed of 3 km/h following a random walk model. The table-based Adaptive Modulation and Coding approach has been employed as given in section 6.1.4.2. Accordingly, the maximum spectral efficiency is $\eta_{\max}$=5 bits/s/Hz. As in Scenario A, users always have data ready to be sent (i.e., full-buffer traffic model), so that each user tries to obtain as much capacity as possible. However, since maximum spectral efficiency has increased with respect to Scenario A, now the target throughput $th_{\text{target}}$ per user is set to 256 kbits/s.

Finally, simulations were carried out either with $N$=12 chunks as in Scenario A or $N$=24 chunks to illustrate that proposed DSA framework works appropriately when the number of chunks is increased (notice that the rare number of possible spectrum assignments in a given scenario is $2^{KN}$, and thus HEUR-DSA or RL-DSA have to select one spectrum assignment among $2^{19\cdot 24}$ possibilities). Other simulation parameters not specified are taken from Table 6.4.

### 6.3.3 Scenario C. Femtocell deployment

In this case, a femtocell scenario without a macrocell layer is used to study the spectrum assignment in a pure decentralized environment. Then, the non-cooperative decentralized DSA framework will be compared with other state-of-the-art strategies.

A total of 10 and 100 circular femtocells are randomly deployed with a uniform distribution in a square area of 500x500 m2 to obtain the performance comparison under two different densities of femtocells. Figure 6.8 illustrates a sample of this scenario for 100 randomly deployed femtocells, where a femtocell radius on 20 m has been taken.

The number of chunks considered is $N$=24 chunks. The macrocell pathloss model is used between femtocells for interference considerations plus an additional wall penetration loss detailed in Table 6.5. Users in a femtocell remain static and the target throughput $th_{\text{target}}$ per user is varied from 128 kbits/s to 2048 kbits/s to simulate different traffic requirements. The number of users per femtocell is randomly selected, being 4 the maximum number of users in any case.

Finally, it is worth to remark that results will be averaged for 100 different femtocell positions and users' distributions. Other simulation parameters are specified in Table 6.5.

**Table 6.4 Simulation parameters for Scenario B**

| | |
|---|---|
| Frame duration | 2 ms |
| Simulated time | 1 hour |
| KPIs averaging time | $l$ =2500 frames |
| Number of cells | $K$ = 19 |
| Cell Radius | $R_c$ = 500 m |
| Number of antennas | 1 Tx, 1 Rx |
| Antenna Patterns | Omnidirectional |
| Maximum Power per chunk | $P_{max}$ = 32.21 dBm |
| Carrier Frequency | 2 GHz |
| Number of Chunks | $N$ =12 or $N$ =24 |
| Chunk bandwidth | 375 kHz |
| Path Loss at d km [dB] | $128.1+37.6\log_{10}(d)$ |
| Path Loss Exponent | $\chi$ = 3.76 |
| Shadowing | Lognormal |
| Standard deviation | $\sigma_s$ =8 dB |
| De-correlation model and distance | [76], 5m |
| Small Scale Fading Model | ITU Ped. A [76] |
| UE thermal noise | -174 dBm/Hz |
| UE noise factor | 9 dB |
| UE speed | 3 km/h |
| UE Satisfaction throughput | $th_{target}$ =256 kbps |
| Number of users | Initially 15 users per cell. From minutes 25 to 35 the number of users per cell is varied (see explanation on text) |
| Scheduling | Proportional Fair |
| Averaging Window size | $T_W$ = 50 frames |
| Adaptive Modulation and Coding | Table-based (see section 6.1.4.2) |

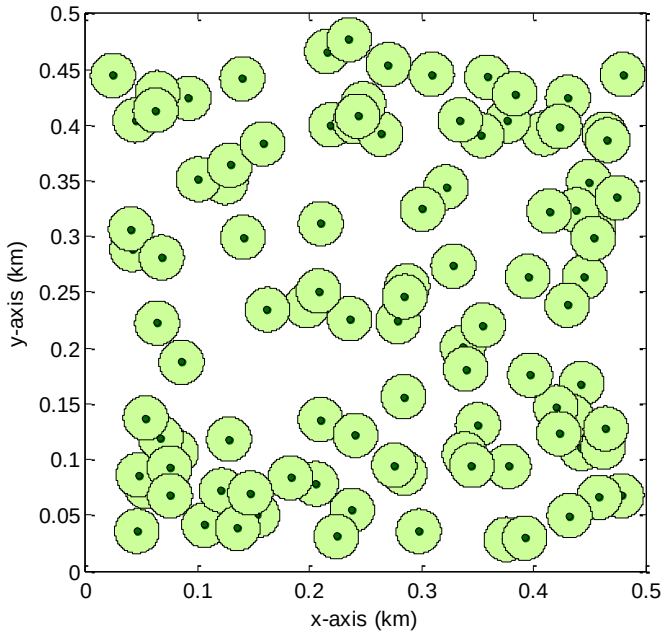


**Figure 6.8 Sample of Scenario C layout with 100 femtocells**

**Table 6.5 Simulation parameters for Scenario C**

| | |
|---|---|
| Frame duration | 2 ms |
| Simulated time | 10 minutes |
| KPIs averaging time | $l$ =2500 frames |
| Number of simulations | 100 |
| Femtocell Radius | 20 m |
| Minimum user distance to Access point | 1 m |
| Antenna Pattern | Omnidirectional |
| Power per chunk | -3.8 dBm (max cell power 10dBm) |
| Carrier Frequency | 2 GHz |
| Number of Chunks | $N$ =24 |
| Chunk bandwidth | 375 kHz |
| Path Loss at $d$ m in dB | $37+32\log_{10}(d)$ |
| Shadowing Standard deviation | $\sigma_s$ =8 dB |
| De-correlation model and distance | [76], 5m |
| Small Scale Fading Model | ITU Ped. A [76] |
| External wall penetration loss | 15 dB |
| UE thermal noise | -174 dBm/Hz |
| UE noise factor | 9 dB |
| UE speed | 0 km/h (static) |
| UE Satisfaction throughput | $th_{\text{target}}$ =128, 512, 1024 or 2048 kbps |
| Number of users | From 1 to 4 randomly selected at femtocell deployment |
| Scheduling | Proportional Fair |
| Averaging Window size | $T_W$ = 50 frames |
| Adaptive Modulation and Coding | See Table 6.2 |

### 6.3.4 Scenario D. Two-layer cellular deployment

This scenario focuses on a two-layer cellular deployment scenario compound of macrocells and femtocells. Co-channel spectrum assignment is considered. The aim is to evaluate the performance of different spectrum assignment strategies. Concretely, special attention is paid to the intercell interference between layers and to the adaptability of the proposed DSA framework, particularly the decentralized approach.

As macrocell deployment, a layout of 7 omnidirectional macrocells is employed, whose configuration values are given in Table 6.6. A total of 15 users per macrocell are uniformly deployed, requiring $th_{\text{target}}$ =256kbits/s per user as satisfaction throughput. Two different subcases have been envisaged for the femtocell layer, where *closed access* is assumed, that is, macrocell users cannot connect to femtocells. This produces worse case interference patterns between macrocell and femtocells.

First, 10 or 100 circular femtocells are randomly positioned in the coverage area of the central macrocell. As an example, Figure 6.9 illustrates a sample scenario with 100 femtocells, where femtocells are represented by light green circles. Configuration values are given in Table 6.6. For both the macrocell and the femtocell layer $N$ =24 chunks are considered. Femtocell users require $th_{\text{target}}$ =512kbits/s.

Second, 12 femtocells inside a building as depicted in Figure 6.10 are considered to evaluate the DSA framework under realistic indoor environment. The building is situated at approximately 200 m of the central macrocell. Each one of the offices has 20x20 m2, and the femtocell is located at the office's center. For indoor coverage, the COST 231 Multi-Wall Model (MWM) is used [77]. Height of the building's walls is 4 m. Inner walls are considered as 'narrow' walls (i.e., with small penetration losses), whereas external walls are considered as 'thick' walls. Penetration losses for doors and windows are also considered. Concretely, penetration losses are given as [external wall, inner wall, door, window]= [15,10,3,1] dB. Besides, two distributions, homogeneous and heterogeneous, of the users in the building are considered. In the homogeneous distribution, 4 users are deployed in each office (i.e., femtocell) whereas in the heterogeneous distribution, half of the offices has 8 users and the rest 4 users. The number of chunks considered in this case for the macrocell and femtocell layer is $N$ =12.

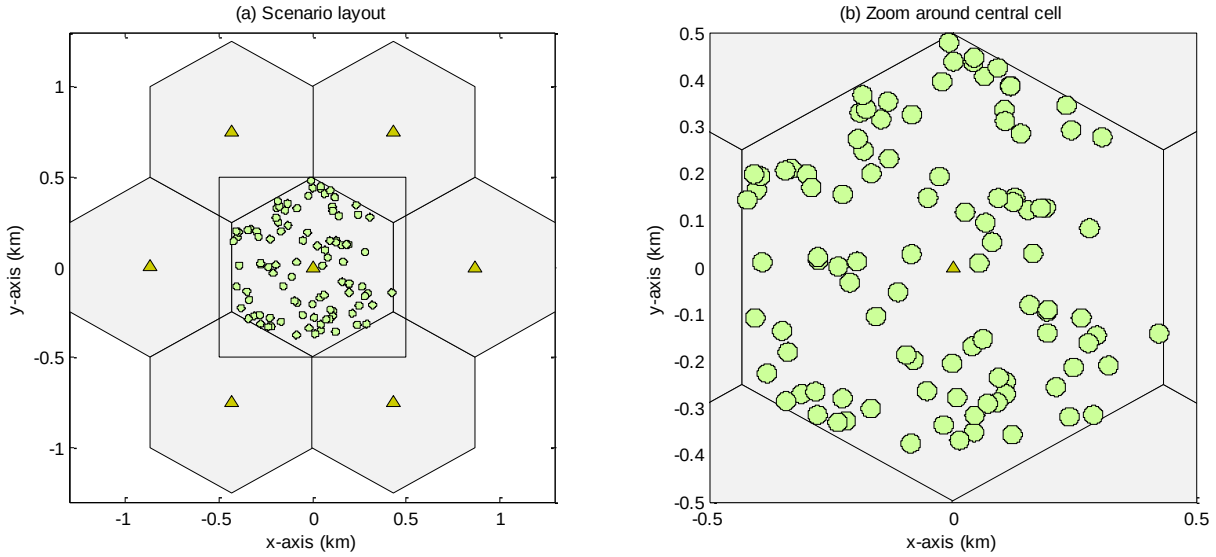


**Figure 6.9 Scenario D with random deployed femtocells.**

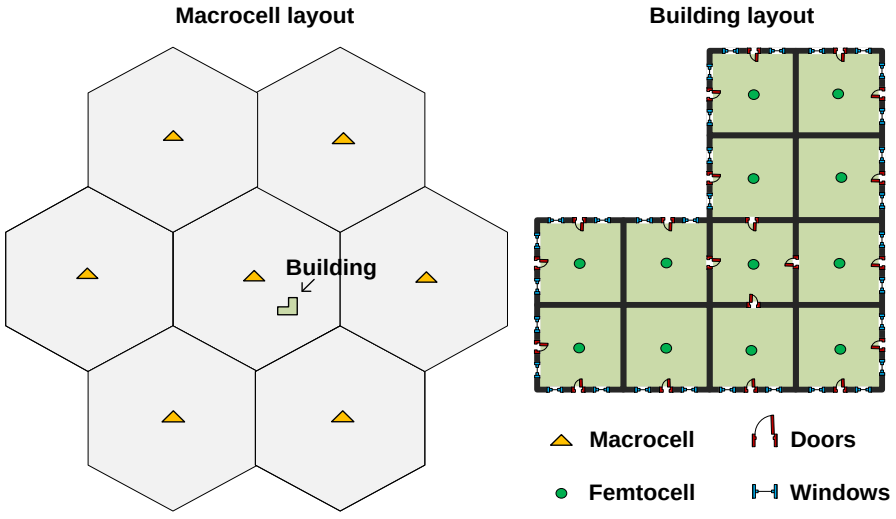


**Figure 6.10 Scenario D layout with building detail including femtocells**

**Table 6.6 Simulation parameters for Scenario D**

| | |
|---|---|
| **Common parameters** | |
| Frame duration | 2 ms |
| Simulated time per simulation | 10 minutes |
| KPIs averaging time | $l$ =2500 frames |
| Number of simulations | 100 |
| Scheduling | Proportional Fair |
| Averaging Window size | $T_W$ = 50 frames |
| Adaptive Modulation and Coding | See Table 6.2 |
| Carrier Frequency | 2 GHz |
| Number of Chunks | $N$ =12 or $N$ =24 |
| Chunk bandwidth | 375 kHz |
| UE thermal noise | -174 dBm/Hz |
| UE noise factor | 9 dB |
| **Macrocell parameters** | |
| Number of cells | $K$ = 7 |
| Cell Radius | $R_c$ = 500 m |
| Number of antennas | 1 Tx, 1 Rx |
| Antenna Patterns | Omnidirectional |
| Maximum Power per chunk | $P_{\max}$ = 43-10log$_{10}$( $N$ ) dBm |
| Path Loss at d km [dB] | 128.1+37.6log$_{10}$(d) |
| Path Loss Exponent | $\chi$ = 3.76 |
| Shadowing | Lognormal |
| Standard deviation | $\sigma_s$ =8 dB |
| De-correlation model | [76] |
| De-correlation distance | 5 m |
| Small Scale Fading Model | ITU Ped. A [76] |
| Number of users | 15 per cell |
| Satisfaction throughput | 256 kbps |
| UE speed | 3 km/h |
| **Femtocell parameters** | |
| Femtocell Radius | 20 m |
| Minimum user distance to Access point | 1 m |
| Antenna Pattern | Omnidirectional |
| Power per chunk | 10-10log$_{10}$( $N$ ) dBm |
| Path Loss at $d$ m in dB | 37+32log$_{10}$($d$) |
| Shadowing | Lognormal |
| Standard deviation | $\sigma_s$ =8 dB |
| De-correlation model | [76] |
| De-correlation distance | 50 m |
| Small Scale Fading Model | ITU Ped. A [76] |
| Penetration losses | [external wall, inner wall, door, window]=[15,10,3,1] dB |
| UE speed | 0 km/h (static) |
| UE Satisfaction throughput | $th_{\text{target}}$ =512, 1024 or 2048 kbps |
| Number of users | - Random deployment: From 1 to 4<br>- Office: 4 or 8 users depending on the distribution |

## 6.4 Summary

This chapter has presented the evaluation methodology that has been followed to obtain the performance results presented in next Chapter 7.

The first part of the chapter was devoted to describe the simulation tool, named OFDMA-DSA Simulator. The purpose of this platform is to compare different spectrum assignment strategies in an OFDMA wireless cellular environment. The simulation tool has been built as a stand-alone application in C/C++. As inputs, it accepts configurable in text files where parameters for cellular deployment, OFDMA, radio propagation, traffic and mobility among others can be configured. As outputs, metrics such as spectral efficiency, throughputs, SINR patterns can be monitored, which are stored in text files for post-processing. Details of the simulator regarding its functional architecture, procedure and implemented models have been given.

The chapter also defined the metrics that will be used to compare the different spectrum assignment strategies. Concretely, as user QoS metric, the dissatisfaction probability has been presented, which defines the probability that average user throughput is below a given target. On the other hand, as network performance metrics, the spectral efficiency in bits/s/Hz and the throughput fairness have been defined. The latter measure is presented as the balance between the throughout obtained by users among the cell area (i.e., between users near and far from the base station). Finally, to measure the capacity of a given strategy to generate spectrum access opportunities for secondary usage the Useful Released Surface has been retained. This metric reveals the amount of spectrum that can be released in a given area for secondary usage.

Finally, the chapter ends with a description of the application scenarios where the proposed spectrum assignment framework has been compared with other state-of-the-art strategies. Four scenarios have been envisaged numbered from Scenario A to Scenario D. Scenario A addresses a macrocell deployment with static distribution of the traffic load per cell where four possibilities from homogeneous distribution to highly heterogeneous distribution have been considered. The objective is to be able to analyze the performance of the spectrum assignment strategies under different spatial distributions of the traffic load. Scenario B follows a similar approach than Scenario A, but the traffic load is changed on-line during simulation to enable the analysis of the dynamics of the DSA framework. Scenario C focuses on a pure decentralized scenario such as a femtocell scenario, whereas Scenario D combines a macrocell and femtocell scenario to illustrate the adaptability of the DSA framework to this particular case.

# 7

# RESULTS

## Outline

# 7 Results

This chapter validates the DSA framework and DSA strategies proposed in this thesis through an exhaustive performance evaluation in the four scenarios presented in Chapter 6, which encompass (*i*) a macrocell deployment to assess the DSA achievements under different static spatial distributions of the traffic load, (*ii*) a macrocell deployment with dynamic variation of the traffic load to insight the dynamism and adaptability of the proposed DSA schemes, (*iii*) a decentralized femtocell deployment to evaluate the suitability of the decentralized DSA scheme in pure distributed cellular scenarios and a combined macrocell, and (*iv*) a femtocell deployment to examine the adaptability of the decentralized DSA schemes in an environment where intercell interference between uncoordinated layers may arise. Then, the centralized and decentralized version of the DSA framework presented in Chapter 4, and the HEUR-DSA and RL-DSA algorithms presented in Chapter 5 will be applied to these layouts to examine their suitability.

Results are obtained by performance comparison with state-of-the-art strategies depending on the scenario. Concretely, Frequency Reuse Factor (FRF), Partial frequency Reuse (PR), and Soft frequency Reuse (SR) strategies presented previously in Chapter 3 have been considered as reference cases for macrocell deployments. These strategies divide the available OFDMA frequency chunks into equal sets that are assigned to cells in a regular basis. In the case of PR and SR an additional division between the central and outer parts of the cell and, accordingly, users therein is employed to motivate a performance improvement for users in the edge part of the cell. On the other hand, the Random-FRF presented also in Chapter 3 has been employed as reference spectrum assignment strategy in femtocell deployments. Under this strategy, the available frequency OFDMA chunks are divided into several subsets, where each femtocell randomly selects one of them to operate. For a high number of subsets, then the probability that two near femtocells interfere each other because of using the same subset is small, but, at the same time the available capacity in the femtocell to satisfy users' requirements is reduced.

Performance comparison is carried out in terms of the key performance indicators (KPIs) defined in Chapter 6 such as spectral efficiency, user's throughput dissatisfaction probability, and throughput fairness. Also the capability of each spectrum assignment strategy to generate opportunities for secondary spectrum usage is studied. To this end, the Useful Release Surface (URS) KPI and spatial spectrum usage maps will be used. Furthermore, the intercell interference between macrocells and femtocells in a two-layer cellular deployment will be examined. Finally, the adaptability and dynamism of the proposed DSA framework based on self-organization is evaluated as well.

In the following, this chapter is organized in two parts devoted to present the results for the centralized and decentralized DSA framework approaches respectively running the HEUR-DSA and RL-DSA algorithms.

## 7.1 Centralized DSA framework

The performance of the HEUR-DSA and RL-DSA algorithms on the centralized approach of the DSA framework is evaluated in this section. Table 7.1 includes the configuration values of the framework, which are common for all the results presented in this section (see Chapter 4 for details).

**Table 7.1 Centralized DSA framework configuration values**

| | |
|---|---|
| DSA algorithm trigger threshold (up side) | $\delta^{up}$ =10% |
| DSA algorithm trigger threshold (down side) | $\delta^{down}$ =0.1% |
| Measurements averaging period | $l$ =5 s |

### 7.1.1 Performance results over static distributions of the traffic load

Performance results presented in this clause are obtained in the Scenario A presented in Chapter 6, which consist of a 19 cell macrocell scenario with static distribution of the traffic load. Concretely, the performance of the spectrum assignment strategies over four different spatial distributions of the traffic load among the cellular deployment is assessed. The number of OFDMA frequency chunks is $N$ =12 chunks. Other simulation details can be found in section 6.3.1 on Chapter 6.

First the performance of the four versions of HEUR-DSA algorithm are exhaustively compared with frequency reuse factors FRF1, FRF3, Partial frequency Reuse (PR) and Soft frequency Reuse (SR). Then, RL-DSA algorithm performance is studied.

#### 7.1.1.1 HEUR-DSA algorithm

The HEUR-DSA configuration for this experiment values are given in Table 7.2 (see Chapter 5 for details). Moreover Table 7.3 contains specific configuration values for the PR, SR, HEUR-DSA3 and also HEUR-DSA4 strategies. Since the chunk assignment policy of HEUR-DSA3 algorithm is analogous to PR scheme, for comparison purposes the maximum number of chunks devoted to each subband in both algorithms is the same. Also, HEUR-DSA4 is configured to have the same inner cell radius as the SR scheme so that they are comparable. The reader is referred to Chapter 3 section 3.1.2 for details.

**Table 7.2 HEUR-DSA configuration values**

| | |
|---|---|
| Initial Margin factor | $\varpi$ = 2.5 |
| Margin Factor Step | $\Delta\varpi$ = 0.05 |
| Margin Factor threshold (up side) | $P_{up}$ =10% |
| Margin Factor threshold (down side) | $P_{down}$ =0.1% |

**Table 7.3 Configuration values for PR, SR, HEUR-DSA3 and HEUR-DSA4 schemes**

| | Number of chunks for central usage | Number of chunks for edge usage | Central Resource Ratio | Inner cell radius | Protection SIR | Max. power per edge chunk | Max. power per central chunk |
|---|---|---|---|---|---|---|---|
| | $C$ | $E$ | $\rho$ | $D$ [m] | $\gamma_0$ [dB] | $P_E$ [dBm] | $P_C$ [dBm] |
| PR | 3 | 9 | 0.5 | 321.52 | 3 | 32.21 | 26.63 |
| SR | 8 | 4 | 0.67 | 371.26 | 3 | 32.21 | 31.33 |
| HEUR-DSA3 | 3 | 9 | 0.5 | 321.52 | 3 | 32.21 | 26.63 |
| HEUR-DSA4 | N/A | N/A | N/A | 371.26 | 3 | 32.21 | 31.33 |

The average number of chunks deployed over the scenario is depicted in Figure 7.1. FRF1, PR, SR and FRF3 schemes manifest no variation in the number of chunks assigned to each cell (12, 6, 12 and 4

chunks respectively out of $N$ =12 chunks available in the system). The average number of chunks per cell for the four HEUR-DSA algorithms is considerably reduced respect to FRF1 or SR and remains around 4 chunks per cell. HEUR-DSA2 algorithm exhibits the lowest values because of its better adaptability to system's requirements by estimating the chunk capacity (contrary to HEUR-DSA1 that maintains the chunk assignment regardless the real chunk capacity).

Dissatisfaction probability is depicted on Figure 7.2. For the homogeneous scenario (see Figure 6.6 where the most loaded cell percentage equals to 5.26%) all HEUR-DSA variants maintain dissatisfaction below target $P_{up}$ =10%. In highly heterogeneous scenarios, HEUR-DSA1, HEUR-DSA2 and HEUR-DSA4 demonstrate up to a 5% of absolute dissatisfaction reduction respect to the best fixed reuse factor. HEUR-DSA3 shows the highest dissatisfaction of the HEUR-DSA algorithms but it remains below PR (up to 12% of absolute dissatisfaction reduction). Thus, HEUR-DSA approaches maintain and even decrease dissatisfaction probability respect to fixed spectrum management schemes.

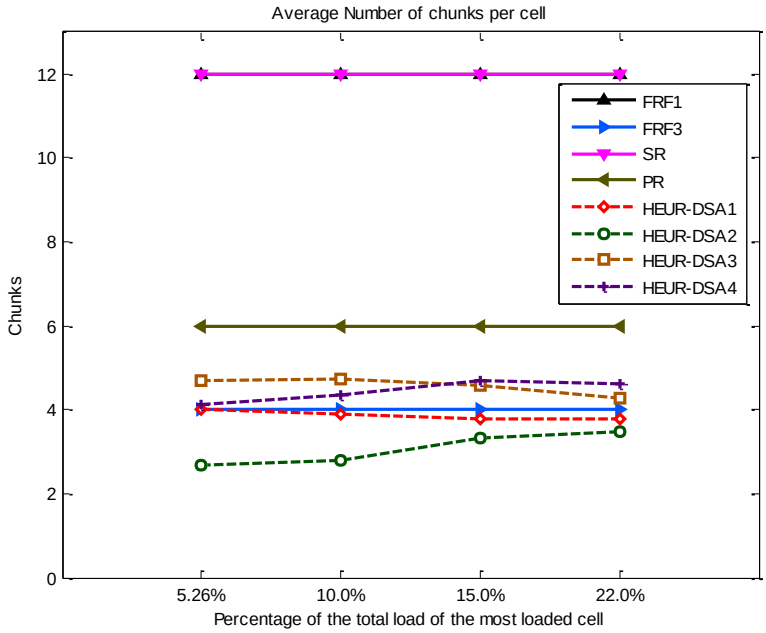


**Figure 7.1 Average number of chunks assigned per cell (FRF1, FRF3, PR, SR, HEUR-DSA1, HEUR-DSA2, HEUR-DSA3 and HEUR-DSA4)**

Average spectral efficiency is represented in Figure 7.3. HEUR-DSA algorithms outperform the fixed schemes with the exception of FRF3 for highly heterogeneous scenarios. However, as shown in Figure 7.2, FRF3 presents the worst users' dissatisfaction for those scenarios. Focusing on the dynamic schemes, HEUR-DSA2 ameliorates in spectral efficiency to HEUR-DSA1 due to its precise adaptation to small variations of the chunk capacity conditions. Both algorithms improve at least 33% the spectral efficiency respect to FRF1 and SR. On the other hand, both HEUR-DSA3 and HEUR-DSA4 algorithms demonstrate slightly poorer spectral efficiency than the other HEUR-DSA algorithms because the minimum number of chunks per cell is two (one for central and another for edge subband) instead of one as it is in the DSA algorithms. Thus, for low loaded cells, HEUR-DSA3 and HEUR-DSA4 algorithms assign more chunks than the minimum required and, as a result, the potential intercell interference increases, avoiding reaching the highest spectral efficiency. In any case, HEUR-DSA3 and HEUR-DSA4 performance is better than PR and SR respectively.

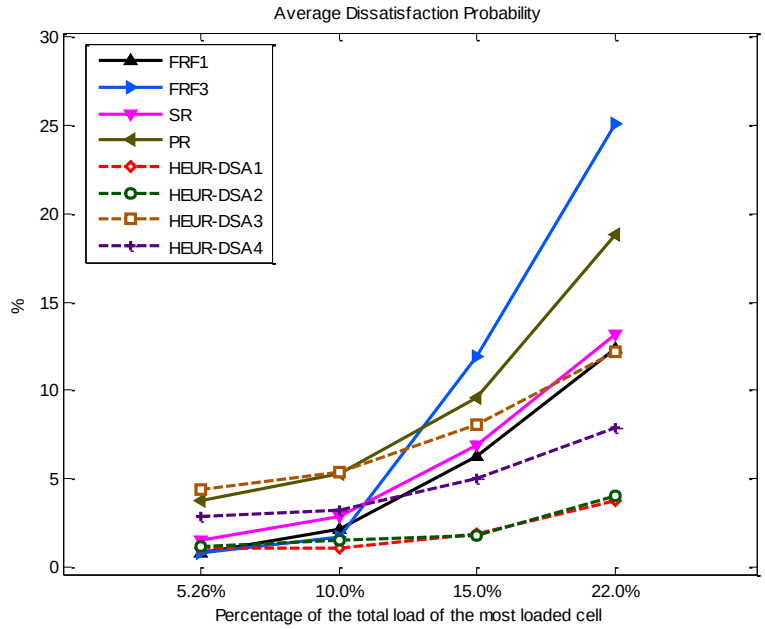


**Figure 7.2 Average dissatisfaction probability (FRF1, FRF3, PR, SR, HEUR-DSA1, HEUR-DSA2, HEUR-DSA3 and HEUR-DSA4)**

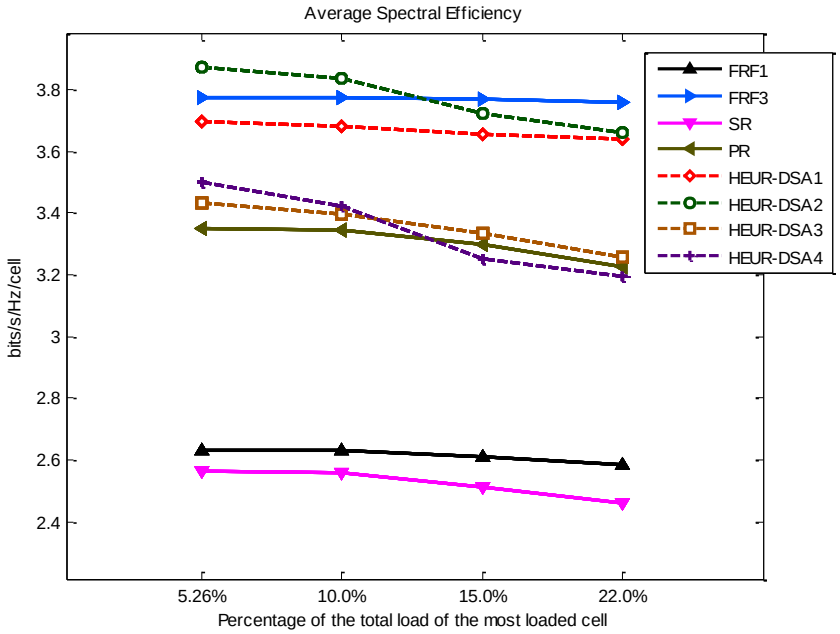


**Figure 7.3 Average spectral efficiency (FRF1, FRF3, PR, SR, HEUR-DSA1, HEUR-DSA2, HEUR-DSA3 and HEUR-DSA4)**

Finally, the Useful Released Area (URS) results are depicted on Figure 7.4. The first tier of cells has been taken as the protection area of primary users (i.e., if a cell is using a chunk, secondary cognitive users cannot use this chunk in the neighboring cells, even if the chunk is free in those cells). Also, the area weighting factor $\omega_a^{(n)}$ is the fraction potential secondary users that would exist in an area assuming that they are uniformly distributed in the scenario. With this definition, the minimum Useful Released Surface value is one chunk free over a cluster of 7 contiguous cells (i.e., 1.70 MHz x Km$^2$ for data in Table 6.3).

Contrary to fixed reuse schemes, HEUR-DSA algorithms evidence that it is possible to release spectrum in regional areas. It can be observed on Figure 7.4 that the URS increases as the users concentrate on a single cell (up to 9 MHz x km$^2$ for HEUR-DSA2). For the homogeneous case, it is impossible for HEUR-DSA1 to release spectrum respecting the protection area. However, the rest of the HEUR-DSA algorithms, by adapting the number of chunks taking into account realistic chunk capacities, enable the spectrum pooling and thus generate opportunities for secondary users.

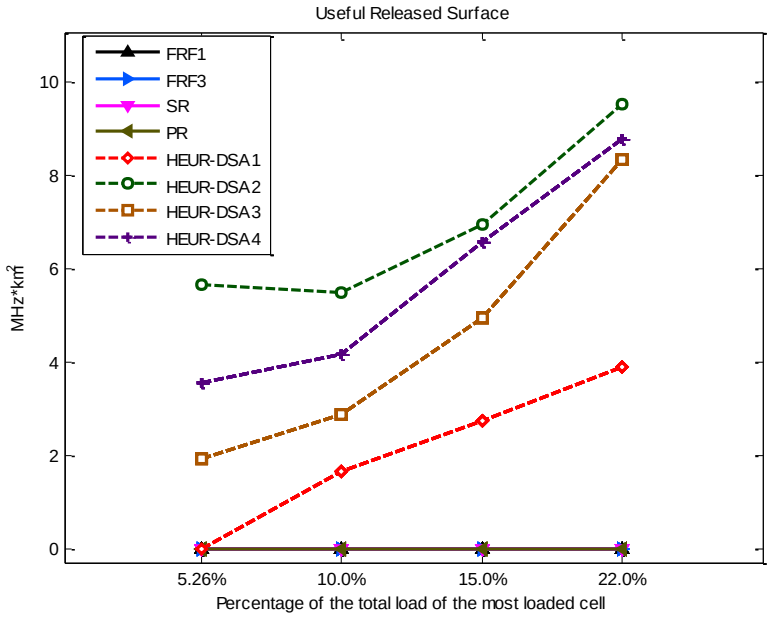


**Figure 7.4 Average Useful Released Surface (FRF1, FRF3, PR, SR, HEUR-DSA1, HEUR-DSA2, HEUR-DSA3 and HEUR-DSA4)**

Some extra features of the HEUR-DSA algorithms appear when metrics are observed per each individual cell. Specifically, cells of interest #3 and #6 highlighted on Figure 6.6 were under study. Figure 7.5 and Figure 7.6 depict individual cell results for CoI#3 and CoI#6 respectively. CoI#3 manifests similar dissatisfaction probability behavior as the global system Figure 7.5(a), since CoI#3 is the dominant cell in the scenario. Spectral efficiency in this cell falls below 2.5 bits/s/Hz for high load (Figure 7.5(b)) because cells around it in Figure 6.6 also increase their load, their number of chunks and thus the intercell interference (CoI#3 almost use all chunks of the system for high loads).

On the other hand, for CoI#6 the user's dissatisfaction maintains below $P_{up}$ =10% (Figure 7.6(a)) as expected. HEUR-DSA2, HEUR-DSA3 and HEUR-DSA4 algorithms slightly increase the dissatisfaction in low loaded cells in order to achieve better spectral efficiency (Figure 7.6(b)) than FRF3 or FRF1, PR and SR respectively. Therefore, DSA algorithms adapt spectrum usage on a cell-by-cell basis and feature an adequate trade-off between spectral efficiency and dissatisfaction probability.

Finally, in order to present the performance of the algorithms that introduce a band division between central and edge subbands, Figure 7.7 depicts the dissatisfaction of the central and edge users for PR, SR, HEUR-DSA3 and HEUR-DSA4 schemes. HEUR-DSA3 and HEUR-DSA4 obtain the best performance for edge especially for high loads, whereas performance for central users is maintained. Concretely around a 20% of absolute dissatisfaction reduction is obtained for HEUR-DSA3 respect to PR in the most loaded case. Therefore, again, dynamic management of the spectrum regarding edge users improves their QoS.

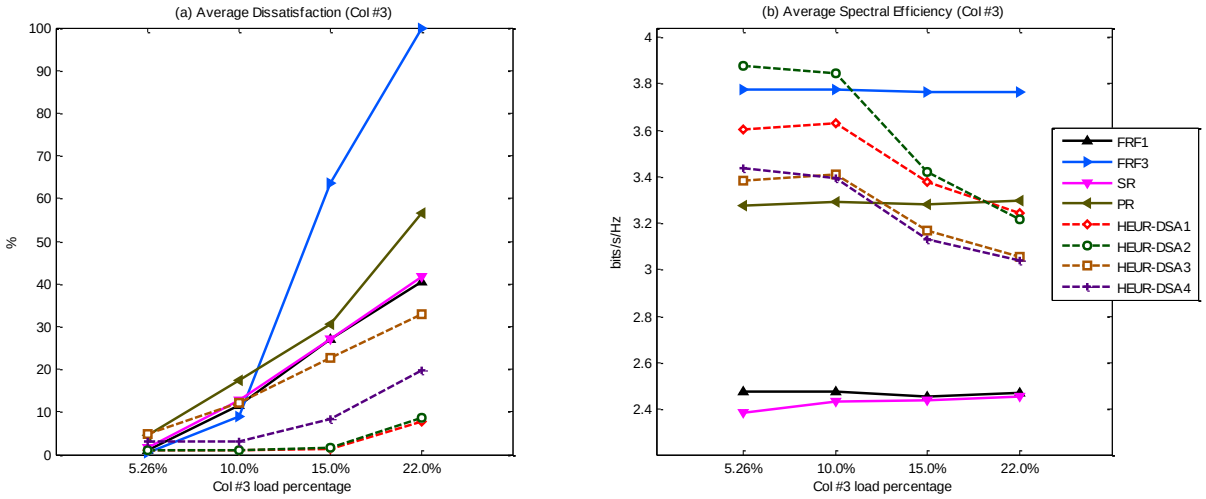


**Figure 7.5 Performance comparison for Cell-of-Interest #3 (FRF1, FRF3, PR, SR, HEUR-DSA1, HEUR-DSA2, HEUR-DSA3 and HEUR-DSA4)**

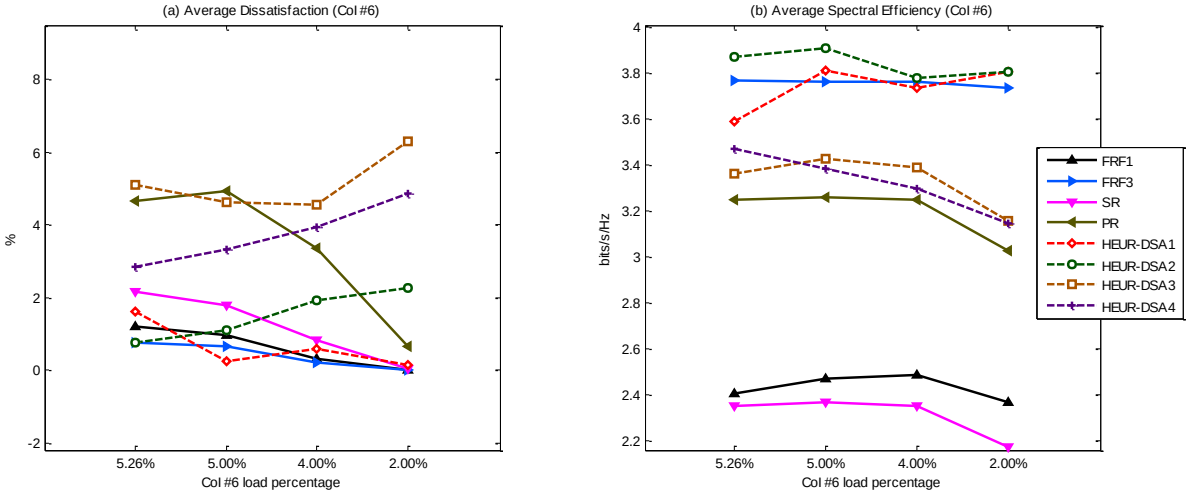


**Figure 7.6 Performance comparison for Cell-of-Interest #6 (FRF1, FRF3, PR, SR, HEUR-DSA1, HEUR-DSA2, HEUR-DSA3 and HEUR-DSA4)**

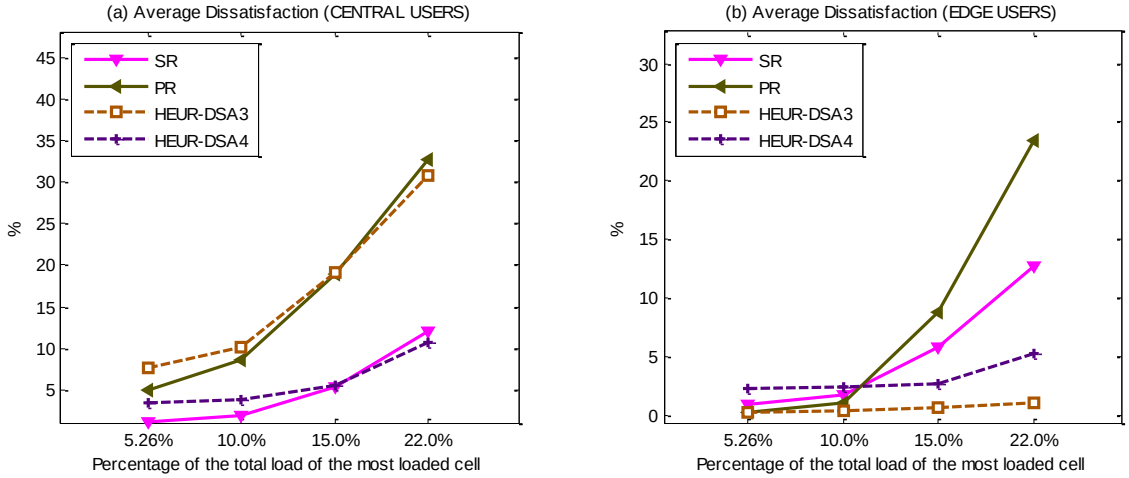


**Figure 7.7 Average dissatisfaction probability for central and edge users (PR, SR, HEUR-DSA3 and HEUR-DSA4)**

Therefore, results above prove that HEUR-DSA approaches improve in practice all the QoS metrics of the primary operator and, in particular, enable the creation of secondary spectrum access opportunities into the licensed band, thus opening a new business opportunities for the primary operator. Finally, HEUR-DSA1 and especially HEUR-DSA2 show better performance than HEUR-DSA3 and HEUR-DSA4. This reflects that the dynamic approaches perform better as the freedom in assigning the chunks into the system band increases. In fact, the division between central and edge subbands incurs a limitation in the way that the chunks are assigned to cells. However, if the network operator decides to deploy PR or SR schemes to control the intercell interference, then it could actually switch to HEUR-DSA3 or HEUR-DSA4 that show better performance than PR or SR respectively.

#### 7.1.1.2 RL-DSA algorithm

The performance of RL-DSA in the centralized DSA framework is compared with FRF1, FRF3, HEUR-DSA1 and HEUR-DSA2 in this subsection. The reward signal given by (5.13) and (5.15) in section 5.2.1.3 in Chapter 5 is used. This reward signal reflects the suitability of a spectrum assignment in terms of user's throughput QoS fulfillment, spectral efficiency and generation of opportunities for secondary spectrum usage. Table 7.4 summarizes the configured values for RL-DSA algorithm.

**Table 7.4 RL-DSA configuration values**

| | |
|---|---|
| Learning rate | $\alpha$ = 100 |
| Learning rate decreasing factor | $\Delta = 10^{-6}$ |
| Reward window averaging factor | $\beta$ = 0.01 |
| Perturbation parameter | $\sigma$ =0.05 |
| Exploratory probability | $p_{\text{explore}}$ =0.1% |
| Maximum number of RL steps | $MAX\_STEPS = 10^6$ |
| Reward upper bound | $R$ =10 |
| Reward Scaling constants | $\lambda$ =1, $\mu$ =0.1 |

Figure 7.8 shows the performance evaluation in terms of spectral efficiency and user's throughout dissatisfaction probability, where the abscissa represents the load of the most loaded cell in each one of the four studied scenarios. Figure 7.8 demonstrates that RL-DSA obtains the best performance in spectral efficiency without compromising the satisfaction of the primary users, which is maintained or even improved respect the fixed spectrum assignment strategies. For instance, in the case of the most heterogeneous load distribution scenario, RL-DSA obtains the best spectral efficiency jointly with FRF3 but absolutely reduces the dissatisfaction of primary users in a 17%.

Furthermore, Figure 7.9(a) shows the average number of chunks per cell in the cellular layout under test. Notice that RL-DSA minimizes the number of chunks per cell but improves spectral efficiency and maintains dissatisfaction probability, as showed in Figure 7.8. Also, Figure 7.9(b) and Figure 7.9(c) show the average number of chunks for the most and least loaded cells respectively. It can be observed that RL-DSA wisely assigns the chunks per cell. For example, HEUR-DSA and RL-DSA increase (decrease) the number of assigned chunks to the most (least) loaded cell. However, RL-DSA organizes the chunks among cells in a way that the satisfaction of the primary users is maintained while the number of chunks is reduced leading to improvements in the spectral efficiency. Also, RL-DSA promotes the generation of spectrum gaps is geographical areas, as shown in the following.

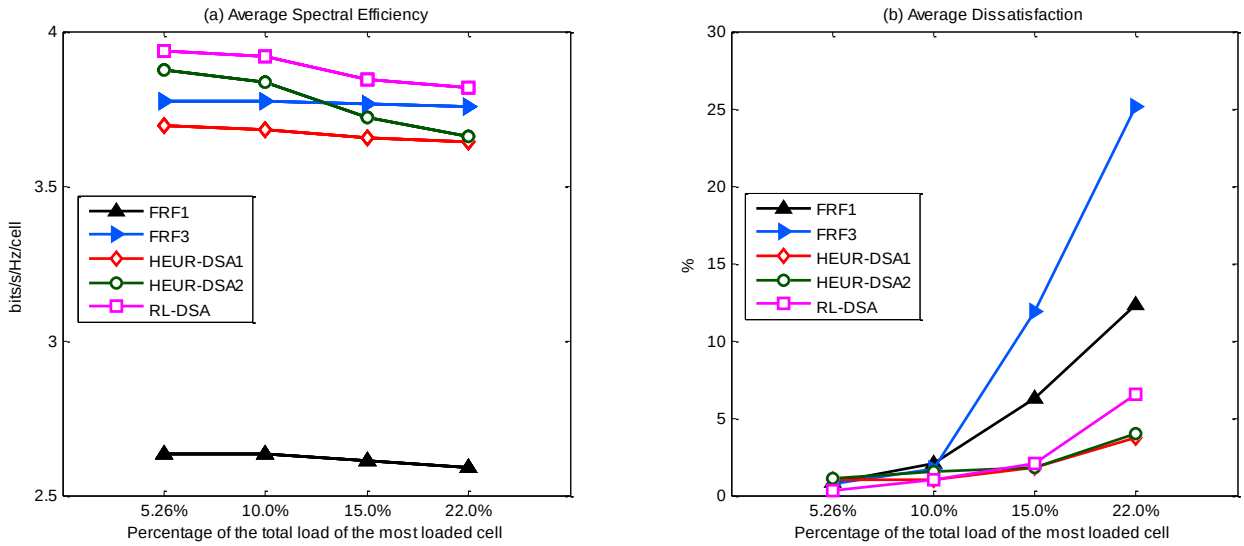


**Figure 7.8 (a) Average spectral efficiency and (b) average dissatisfaction probability (FRF1, FRF3, HEUR-DSA1, HEUR-DSA2 and RL-DSA)**

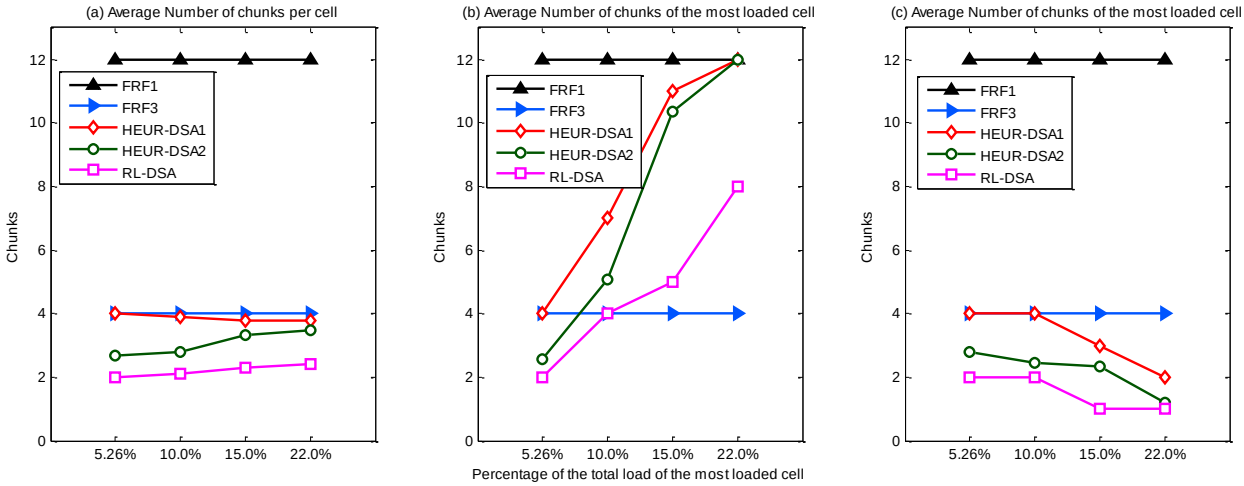


**Figure 7.9 Average number of chunks per cell (FRF1, FRF3, HEUR-DSA1, HEUR-DSA2 and RL-DSA)**

Figure 7.10 shows the great capacity that RL-DSA exhibits for releasing spectrum for secondary access. For instance, at least 10 MHz*Km$^2$ of improvement in Useful Released Surface (URS) can be obtained respect the rest of studied strategies. Since URS takes into account the chunks that are not used in a cell and in cells around it, this means that there are areas in the scenario where a chunk or a group of them can be released from the primary spectrum to be used for a secondary market, without causing harmful interference to primary communications. As an example, Figure 7.11 illustrates the number of chunks per cell that are suitable for secondary spectrum usage in that cell (i.e., chunks that are not assigned to a cell and its neighbors). The figure compares HEUR-DSA2 and RL-DSA for the case where traffic load high heterogeneously distributed (i.e., most loaded cell load equal to 22%). Clearly, around

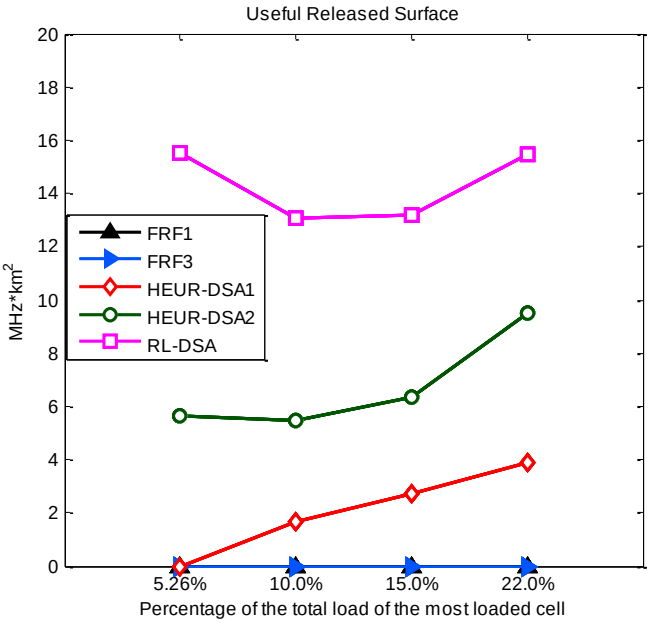


**Figure 7.10 Average Useful Released Surface (FRF1, FRF3, HEUR-DSA1, HEUR-DSA2 and RL-DSA)**

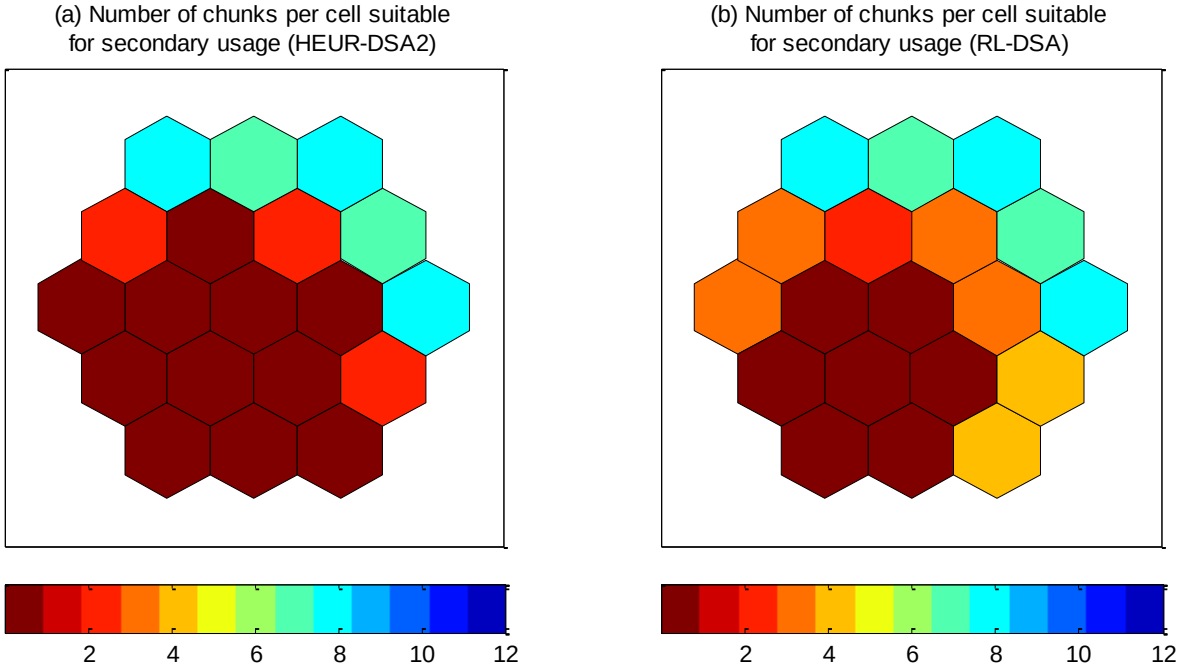


**Figure 7.11 Number of chunks per cell suitable for secondary usage. (a) HEUR-DSA2. (b) RL-DSA**

the most loaded cell there are no opportunities for secondary spectrum usage because primary traffic load prevents so. However, in the set of cells where the traffic load is decrease, RL-DSA outperforms HEUR-DSA2, and up to 8 chunks out of 12 could be, for instance, leased to a secondary spectrum market.

Finally, Figure 7.12 and Figure 7.13 depict the cells with chunks suitable for secondary usage for HEUR-DSA2 and RL-DSA. In the case of RL-DSA the same chunk could be released in a vast area of contiguous macrocells. For instance, chunk #5 can be released in 11 contiguous cells. On the other hand, HEUR-DSA2 only released a chunk in 7 contiguous cells at most. In any case, these results improve fixed FRF strategies, where these secondary usage opportunities are not possible.

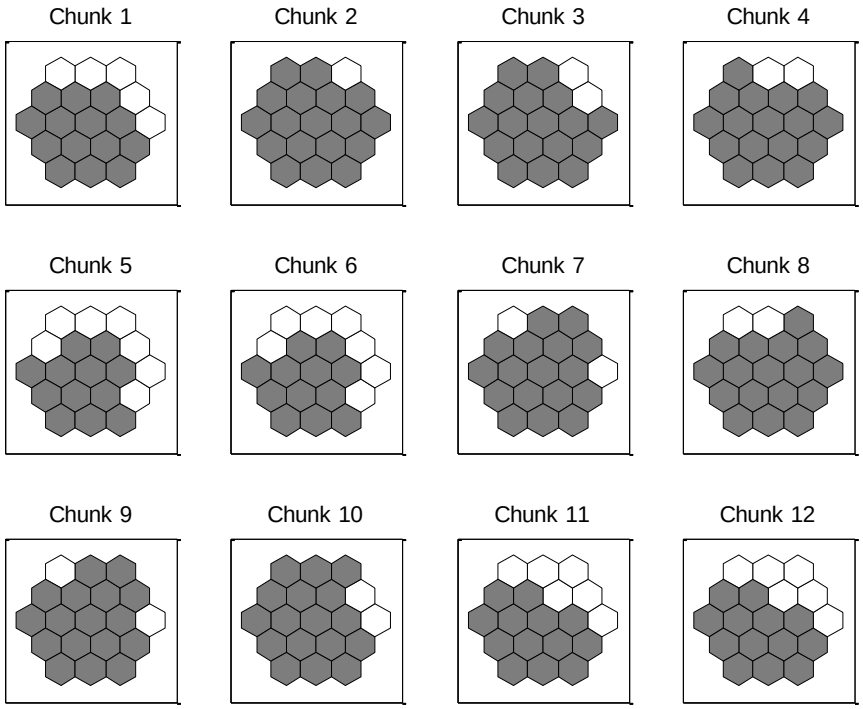


**Figure 7.12 HEUR-DSA2: cells with chunks suitable for secondary spectrum usage (in white).**

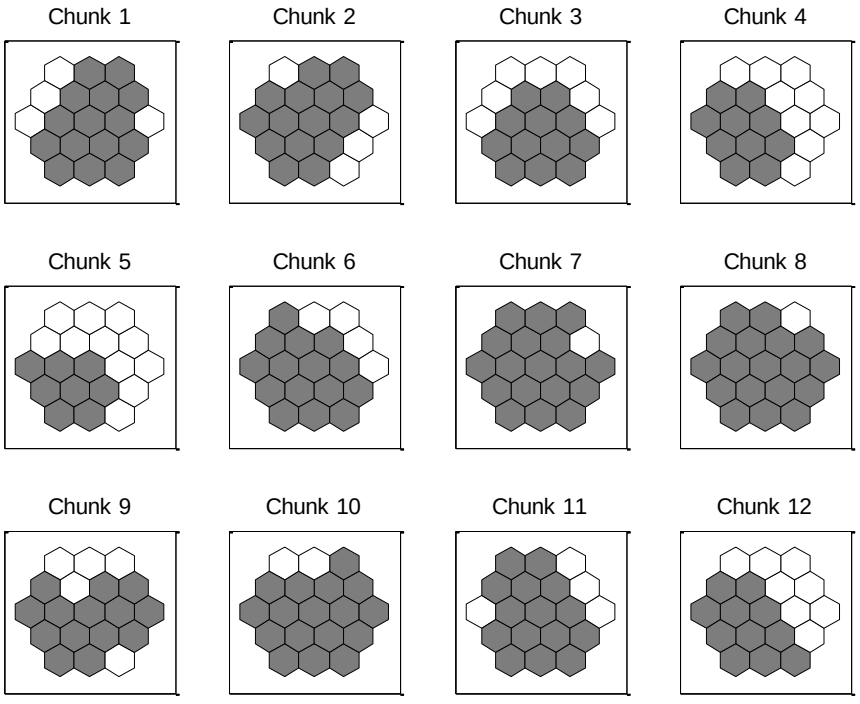


**Figure 7.13 RL-DSA: cells with chunks suitable for secondary spectrum usage (in white).**

### 7.1.2 Performance results over dynamic distribution of the traffic load

Results presented in this section are obtained in Scenario B presented in Chapter 6, where a traffic load is varied during simulation to show the adaptability of the DSA framework and DSA algorithms. The scenario is composed of 19 macrocells and $N$ = 12 chunks are available. The performance of FRF1, FRF3, PR, SR, HEUR-DSA2 and RL-DSA is simulated during 1 hour, where from minute 25 to minute 35 the traffic load is varied from a homogeneous distribution to a heterogeneous distribution according to different traffic load variation rates per cell. Concretely, results are presented for $u_{type1}$=2, $u_{type2}$=1 and $u_{type3}$=-, where $u_{typex}$ stands for the variation rate in users per minute for cells of type $x$ (see section 6.3.2 in Chapter 6 for details). Simulation parameters can be found in Table 6.4 in Chapter 6 as well. Also, configuration values for PR, SR, HEUR-DSA2 and RL-DSA for this experiment are given in Table 7.5.

**Table 7.5 Configuration values for simulations over dynamic traffic load distribution**

| Partial frequency Reuse (PR) | |
|---|---|
| Number of chunks of central cell | $C$ =6 |
| Number of chunks of edge cell | $E$ /3=2 |
| Centre-Edge pathloss threshold | -100 dBm |
| **Soft frequency Reuse (SR)** | |
| Number of chunks of central cell | $C$ =8 |
| Number of chunks of edge cell | $E$ = 4 |
| Centre-Edge pathloss threshold | -100 dBm |
| **HEUR-DSA2** | |
| Initial Margin factor | $\varpi$ = 1.5 |
| Margin Factor Step | $\Delta\varpi$ = 0.15 |
| Margin Factor threshold (up side) | $P_{up}$ =10% |
| Margin Factor threshold (down side) | $P_{down}$ =0.1% |
| **RL-DSA** | |
| Learning rate | $\alpha$ = 100 |
| Learning rate decreasing factor | $\Delta$ =$10^{-6}$ |
| Reward window averaging factor | $\beta$ = 0.01 |
| Perturbation parameter | $\sigma$ =0.05 |
| Exploratory probability | $p_{\text{explore}}$ =0.1% |
| Maximum number of RL steps | $MAX\_STEPS$ =$10^6$ |
| Reward upper bound | $R$ =10 |
| Reward Scaling constants | $\lambda$ =1, $\mu$ =0.1 |

#### 7.1.2.1 Spectral efficiency and dissatisfaction probability

Figure 7.14 depicts the average dissatisfaction probability and spectral efficiency evolution during simulation respectively. As shown in previous section RL-DSA improves the average dissatisfaction probability with respect to the fixed spectrum assignment strategies and shows similar behavior with respect to HEUR-DSA2. Dynamic strategies provide dramatic improvements with respect to fixed strategies especially when the distribution of the load is heterogeneous (i.e., for time greater than 35 minutes). For instance, in this case the dissatisfaction probability is more than four times lower than that for FRF3.

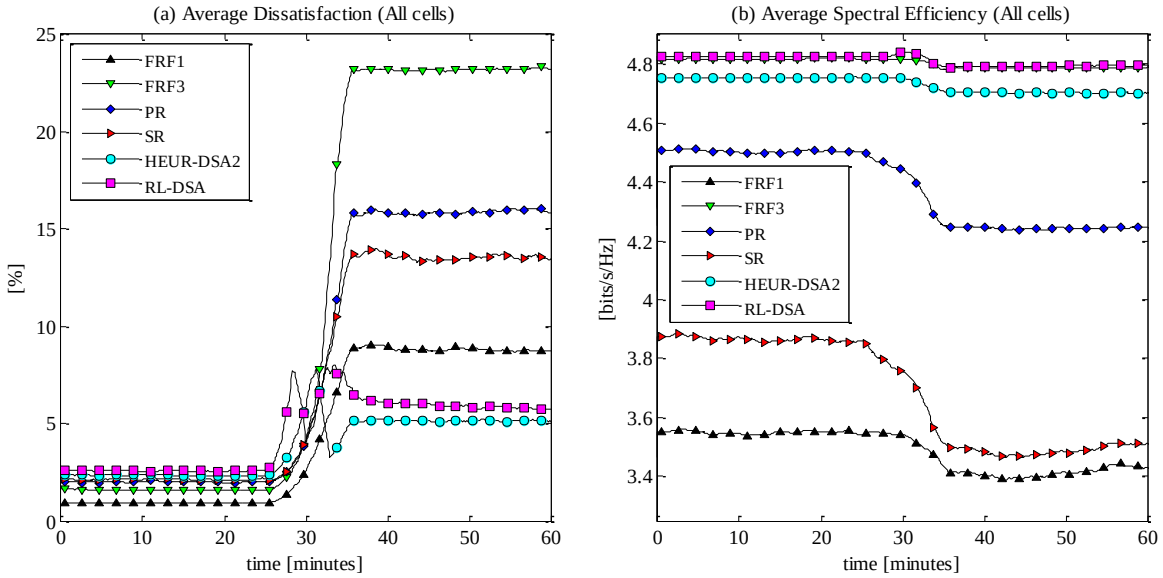


**Figure 7.14 Performance comparison in dynamic traffic load scenario (FRF1, FRF3, PR, SR, HEUR-DSA2, RL-DSA). (a) Average Dissatisfaction Probability. (b) Average Spectral Efficiency**

Moreover, Figure 7.15 and Figure 7.16 depict the average dissatisfaction probability and the spectral efficiency evolution for each type of cell respectively. Results for cells #1, #3 and #9 in Figure 6.7 are presented, as representative ones for type of cells 1, 2 and 3 respectively (analogous results were obtained for other cells of the same type). Notice that RL-DSA maintains a dissatisfaction probability below 10% target for all type of cells and attains the best spectral efficiency for all cells. Thus, RL-DSA achieves the best trade-off between users' satisfaction and spectral efficiency.

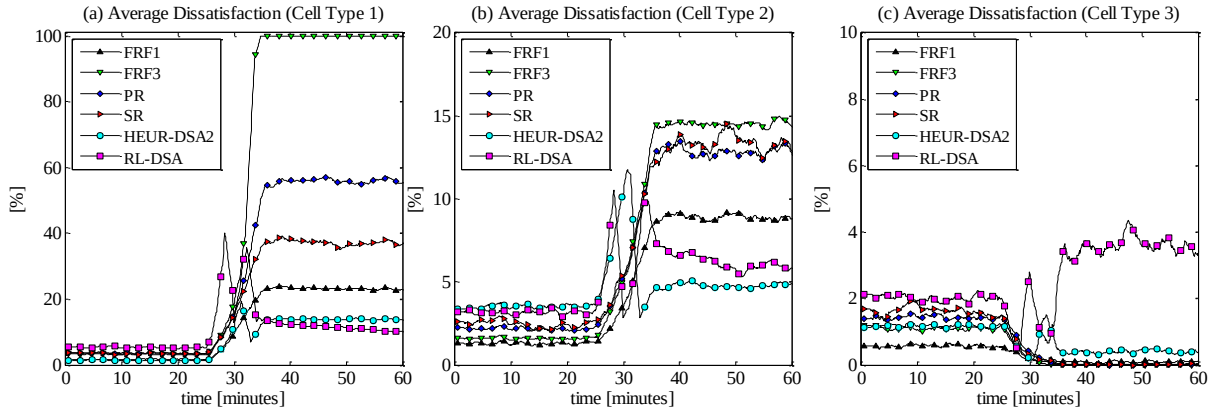


**Figure 7.15 Average dissatisfaction probability per cell comparison in dynamic traffic load scenario**

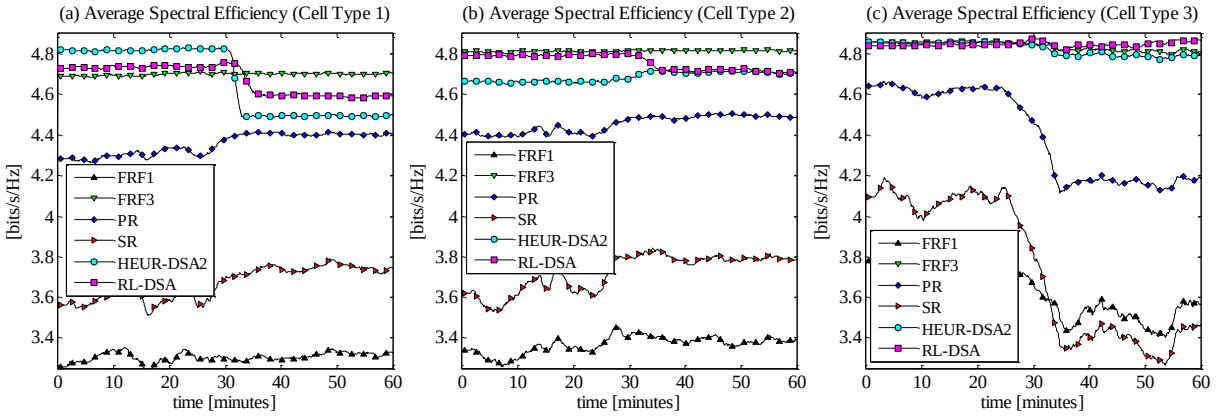


**Figure 7.16 Average spectral efficiency per cell comparison in a dynamic traffic load scenario**

Regarding the adaptability and dynamism of the DSA framework, Figure 7.17 and Figure 7.18 show the evolution of the dissatisfaction probability and assigned chunks per cell for RL-DSA. It can be observed that between minutes 25 and 35 there is an increment of the dissatisfaction probability due to the traffic load variation. Each time that the dissatisfaction probability surpass $\delta^{up}$ =10% (marked in Figure 7.17 with arrows) RL-DSA is triggered by the Status Observer in the centralized DSA Controller. As a result, the number of chunks per cell is varied, being increased in cells of type 1 (i.e., cell #1) and type 2 (i.e., cells #2, #3, #4, #5, #6 and #7), and decreased in the rest of the cells. This supposes a drastic reduction of the dissatisfaction probability.

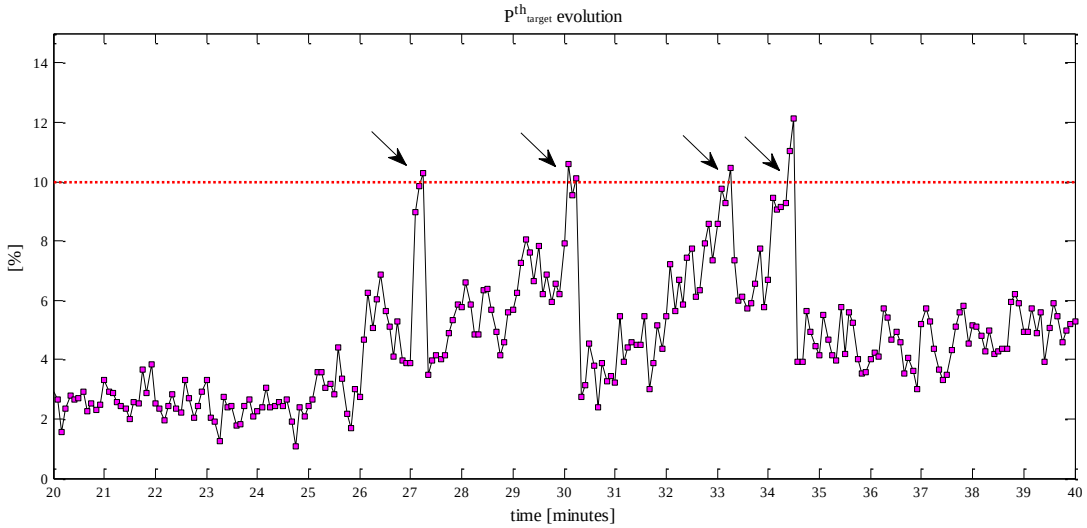


**Figure 7.17 Illustration of the DSA framework dynamism. Dissatisfaction probability evolution**

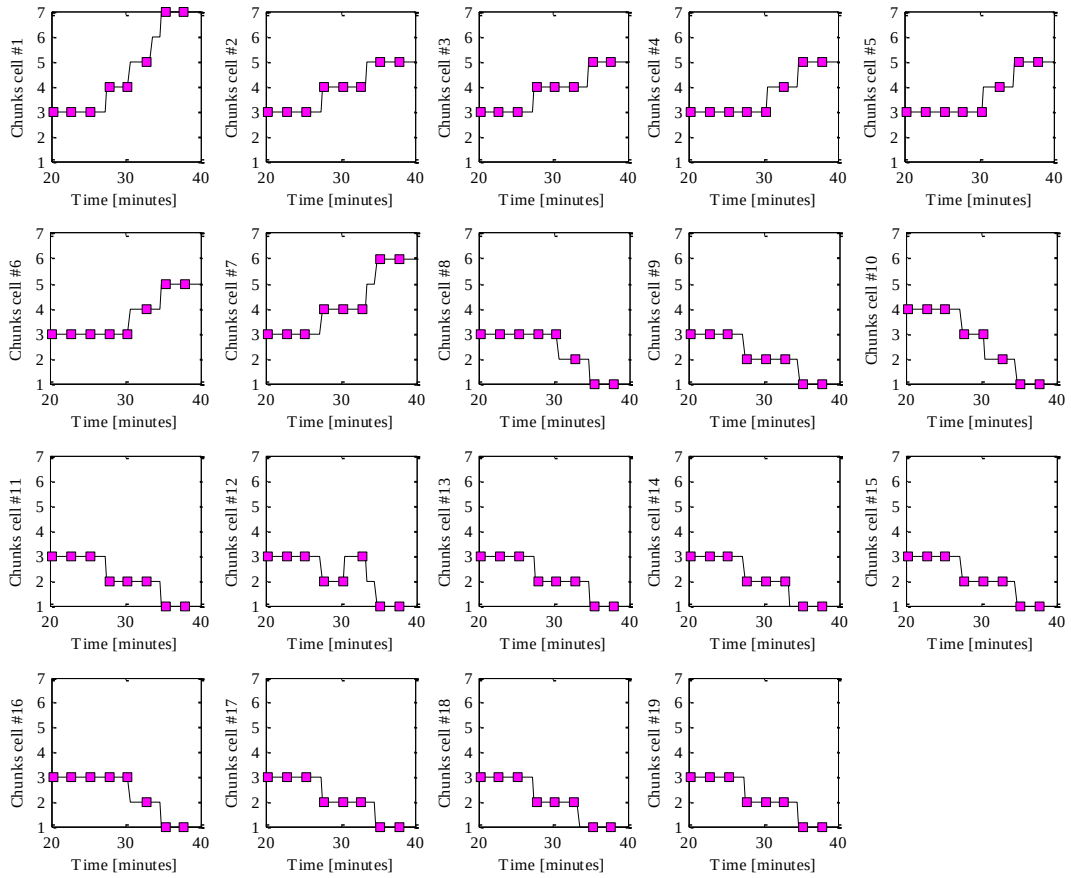


**Figure 7.18. Illustration of the DSA framework dynamism. Chunk assignment per cell evolution**

### 7.1.2.2 Throughput Fairness

Throughput fairness performance is studied in Figure 7.19, which shows the throughput cumulative distribution function normalized to average throughput for all studied schemes at points 20 minutes and 50 minutes (i.e., when a homogeneous or heterogeneous distribution of the load is deployed over the scenario, respectively). HEUR-DSA2 and RL-DSA approach achieve good results in fairness fulfillment, especially for a heterogeneous distribution of the traffic load. On the other hand, FRF3 and SR violate the Normalized Throughput Bound for a heterogeneous distribution of the traffic load.

Analyzing the 5 percentile (i.e., the normalized throughput obtained by at least 95% of the users), RL-DSA is the scheme that best balances system throughput among users, where 95% of users obtain around 75% and 65% of average throughput at points 20 minutes and 50 minutes respectively. Particularly, the evolution of the 5th percentile is depicted in Figure 7.20, where it is shown that RL-DSA is always above the rest of strategies.

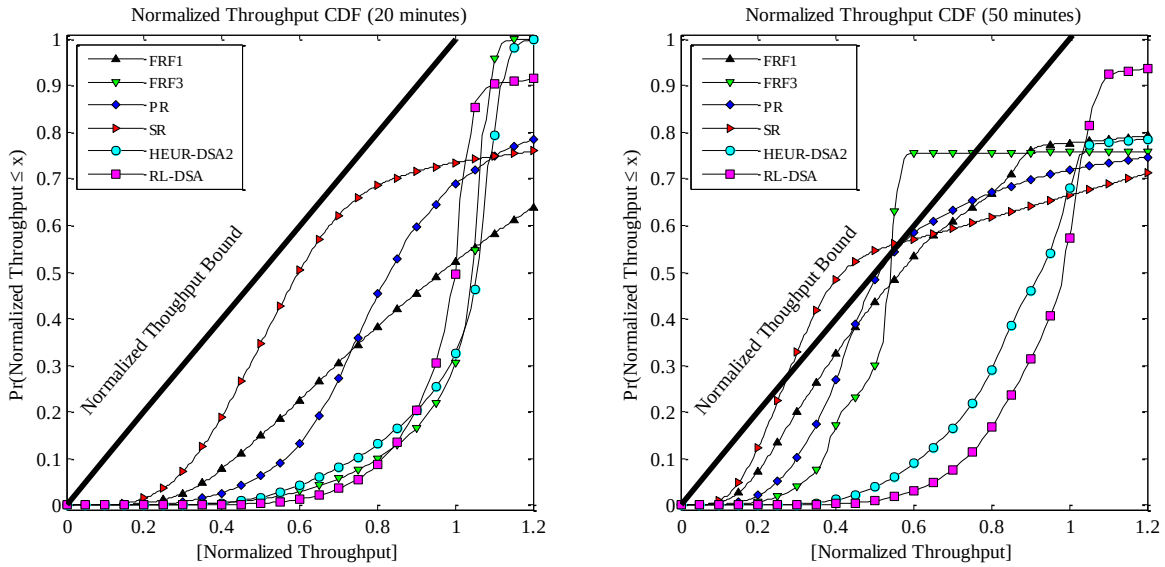


**Figure 7.19 Throughput fairness performance comparison. (a) Fairness at 20 minutes. (b) Fairness at 50 minutes.**

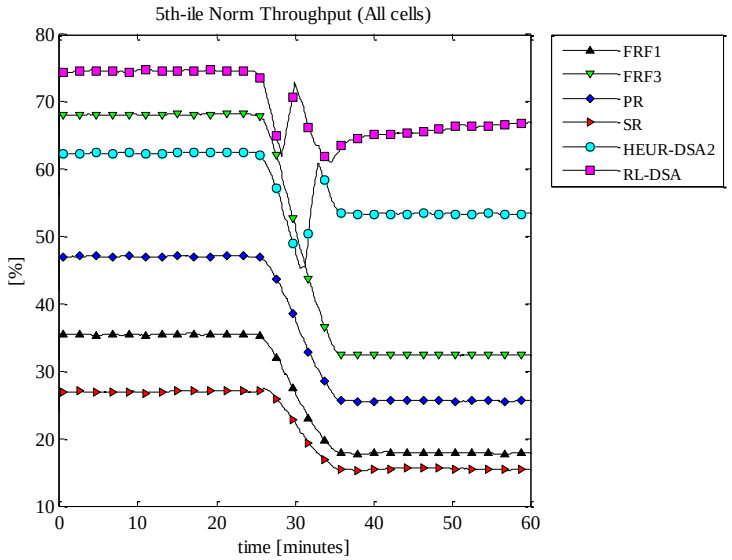


**Figure 7.20 Evolution of 5-th percentile of the normalized throughput**

### 7.1.2.3 Secondary spectrum usage opportunities

Figure 7.21 depicts the chunk usage per cell. Specifically, Figure 7.21(a) shows the average number of non-used chunks per cell in the scenario demonstrating that RL-DSA is the strategy that leaves more free chunks. However, it is also interesting to see how these non-used chunks are distributed. Figure 7.21(b) shows the average number of non-used chunks in clusters of adjacent cells in the scenario. That is, chunks that are not used in a cell and all its adjacent cells and thus would be more appropriate for secondary usage since secondary transmissions would not cause interference in a wider region.

Observe that RL-DSA is the unique strategy that generates these spectrum usage opportunities during the complete simulation. Therefore, RL-DSA assigns the right amount of spectrum per cell so that users obtain the satisfaction throughput, but not more.

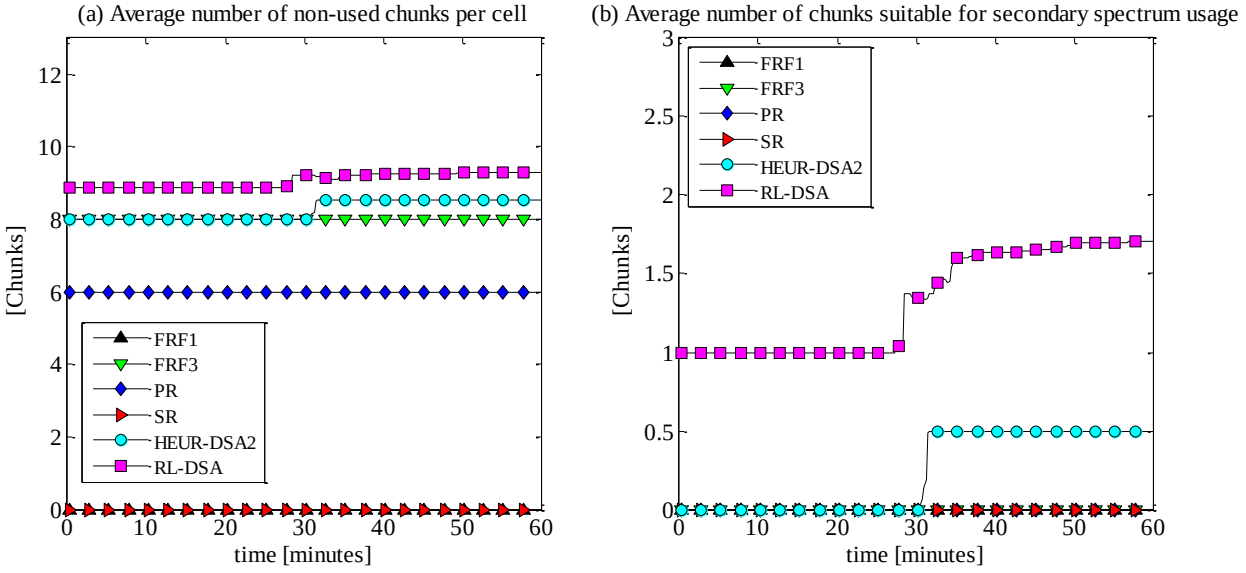


**Figure 7.21 (a) Average number of non-used chunks per cell. (b) Average number of non-used chunks in clusters of adjacent cells (suitable for secondary usage).**

Finally, Figure 7.22 shows the distribution of chunks suitable for secondary usage in the homogeneous and heterogeneous distribution of the traffic load. In the homogeneous distribution, there are less opportunities because RL-DSA tends to distribute chunks avoiding that contiguous cells use the same chunk and thus mitigating intercell interference. However, in the heterogeneous distribution, clearly, the traffic load in central cells limits the number of chunks free.

## 7.2 Decentralized DSA framework

The decentralized version of the DSA framework is evaluated in this part of the chapter. Table 7.6 includes the configuration values of the framework, which are common for all the results presented in this section (see Chapter 4 for details).

**Table 7.6 Decentralized DSA framework configuration values**

| | |
|---|---|
| DSA Algorithm execution period per cell | $L$ =60 s |
| Measurements averaging period | $l$ =5 s |

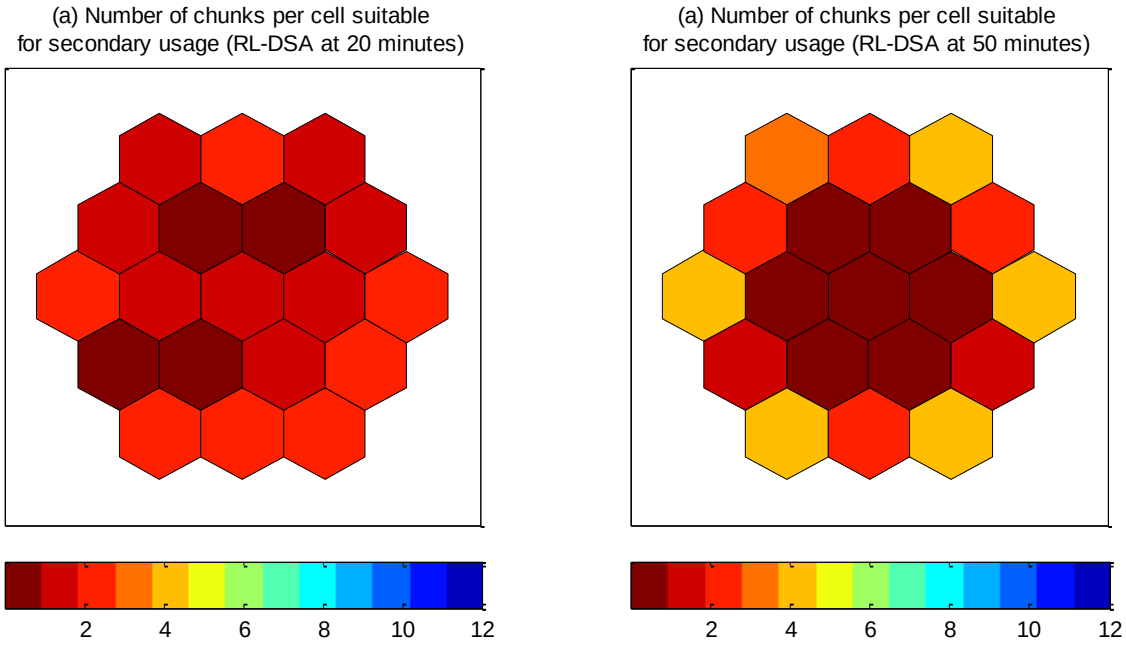


**Figure 7.22 Number of chunks per cell suitable for secondary usage.**

## 7.2.1 Application to a macrocell deployment

The scenario B used in section 7.1.2 is employed here, where the number of available chunks has been doubled (i.e., $N$ =24 chunks) to show that DSA framework operates appropriately with higher number of chunks. Also the traffic load variation rates per cell are set to $u_{type1}$ =6, $u_{type2}$ =0 and $u_{type3}$ =-1 to test highly heterogeneous distributions of the traffic load.

### 7.2.1.1 Performance comparison

Performance comparison is done between FRF1, FRF3, PR, the centralized DSA framework analyzed in previous sections and the decentralized framework for both cooperative and non-cooperative schemes (see Chapter 4 for details). As DSA algorithm, RL-DSA has been employed, where configuration values have been taken from. In this case, the reward signal per cell selected is given by (5.14) in Chapter 5. This reward signal recompenses the SINR fulfillment in a cell for a given spectrum assignment.

Figure 7.23 shows a spectral efficiency comparison between considered schemes. It is clear that RL-DSA strategies overcome the performance attained by the reuse factor strategies. Logically, the centralized approach achieves the best performance but the distributed approaches presented here demonstrate a very close spectral efficiency performance. Furthermore, RL-DSA strategies show a very satisfactory behavior in terms of users' QoS compliance. Concretely, Figure 7.24 shows the average dissatisfaction probability. As it can be seen in the figure, reuse schemes are not adapted to heterogeneous distribution of the load (from minute 35) and thus they obtain poor performance in dissatisfaction probability. In contrast, RL-DSA maintains a reduced dissatisfaction in both homogeneous and heterogeneous spatial distribution of the traffic load.

It is worth noticing the close performance shown by the decentralized strategies in relation to the centralized strategy. This certainly favors the usage of the decentralized approaches because they also add robustness against failure of either a given cell or the centralized controller in charge of executing the centralized algorithm. Furthermore, comparing the cooperative and the non-cooperative schemes, the

cooperative spectrum assignment scheme shows slightly better performance than the non-cooperative scheme because a given cell handles with the precise spectrum assignment in adjacent cells at the moment of RL-DSA execution.

On the other hand, the non-cooperative scheme relies on measurements to estimate the spectrum usage of adjacent cells. These measurements need to be averaged during a certain period of time (i.e., the measurements averaging period). In case that the spectrum assignment in adjacent cells changes during the measurements averaging period, the estimation of the spectrum usage in adjacent cells could be less accurate. However, thanks to the random selection of the initial execution time, the long RL-DSA execution period ( $L$ ), and a sufficiently short measurements averaging period (i.e., $l \ll L$ ), results show that the non-cooperative scheme attains similar performance to the cooperative scheme. Moreover, the non-cooperative spectrum assignment scheme avoids the need of signaling between cells. Hence it raises a good choice in deployments where the intercell signaling interface is not available as in femtocell deployments.

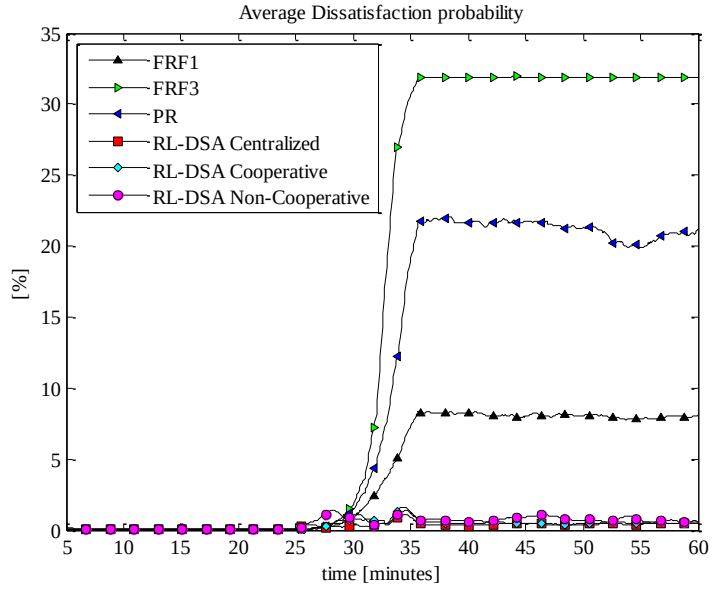


**Figure 7.23 Average dissatisfaction probability comparison (macrocell scenario).**

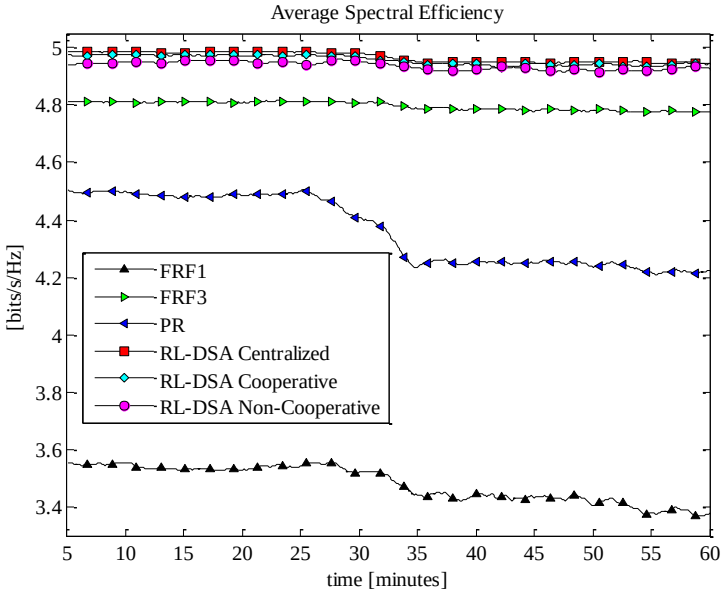


**Figure 7.24 Average spectral efficiency comparison (macrocell scenario).**

Finally, Figure 7.25 shows the average SINR Cumulative Distribution Function (CDF) comparison. RL-DSA strategies exhibit better performance than FRF strategies, what certainly allows the cellular network to achieve the good tradeoff between spectral efficiency and a QoS compliance shown above.

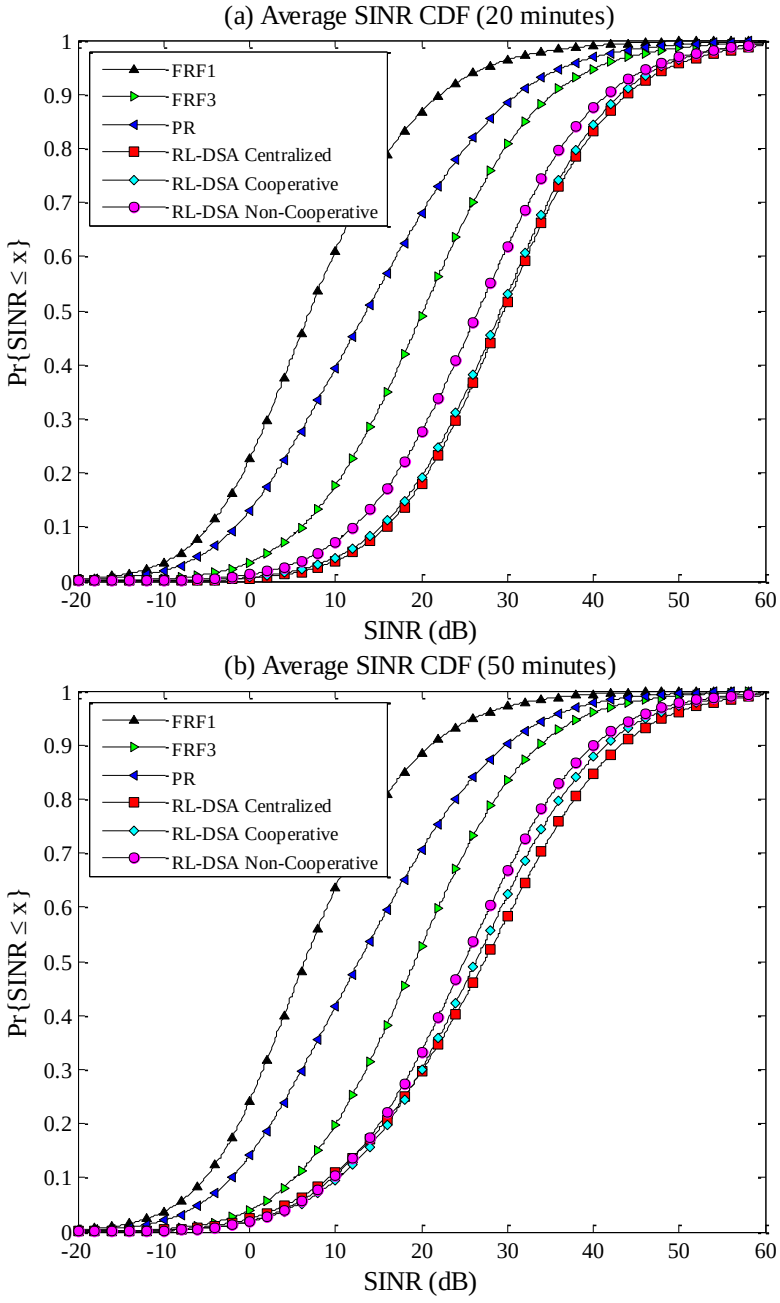


**Figure 7.25 Average SINR Cumulative Distribution Function (CDF) comparison (macrocell scenario).**

### 7.2.1.2 Self configuration of new deployed cell sites

This case study pretends to demonstrate, through an illustrative example, the adaptability of the proposed decentralized self-organized DSA framework with RL-DSA. Suppose that a traffic hot-spot emerges on the macrocell scenario as shown in Figure 7.26. Traffic hot-spot has 30 uniformly distributed static users that connect to nearest macrocell and has a radius of 100 meters. In addition to that, there are

15 uniformly distributed users already operating in each cell area. Then, the performance in most affected macrocells (average dissatisfaction probability and spectral efficiency) is negatively impacted by the hot-spot, whose users experience low signal strength from macrocells, and hence are more sensitive to intercell interference. To cope with this loss of performance, a microcell is activated in the hot-spot area at a certain point of the time *T*. Then, users in the hot-spot perform handover to the microcell after microcell's activation.

Figure 7.27 shows the adaptability exhibit by the microcell and macrocells 1, 5, and 6 in terms of spectrum rearrangement and performance improvement. Before the activation of the microcell, macrocells use between 3 and 5 chunks and their average dissatisfaction probability and spectral efficiency are poor, since they cannot cope with the heavy traffic load in the hot-spot. After the activation of the microcell, spectrum is dynamically managed in micro- and macrocells, activating 4 chunks in the microcell and only 2 chunks per macrocell, which are enough to cope with macrocell users' requirements out of the coverage area of the microcell. This spectrum rearrangement is achieved in two executions of the decentralized framework, whose execution period per cell is of $L$=1 minute. Also, performance in terms of dissatisfaction probability and spectral efficiency is considerably improved. Particularly, it can be seen that dissatisfaction improves very significantly by falling below 5% and spectral efficiency increases accordingly after microcell activation.

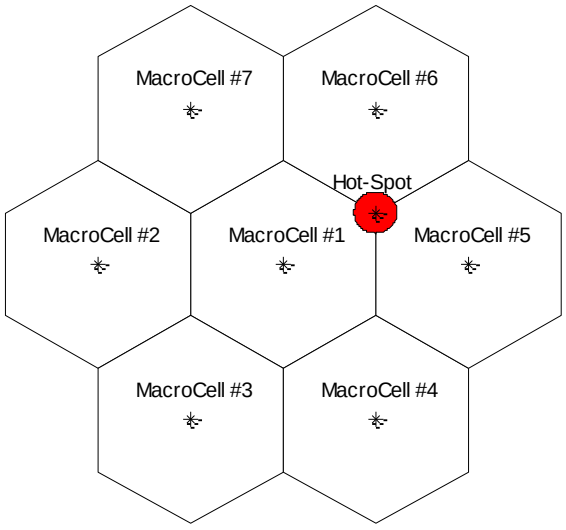


**Figure 7.26 Scenario layout with hot-spot**

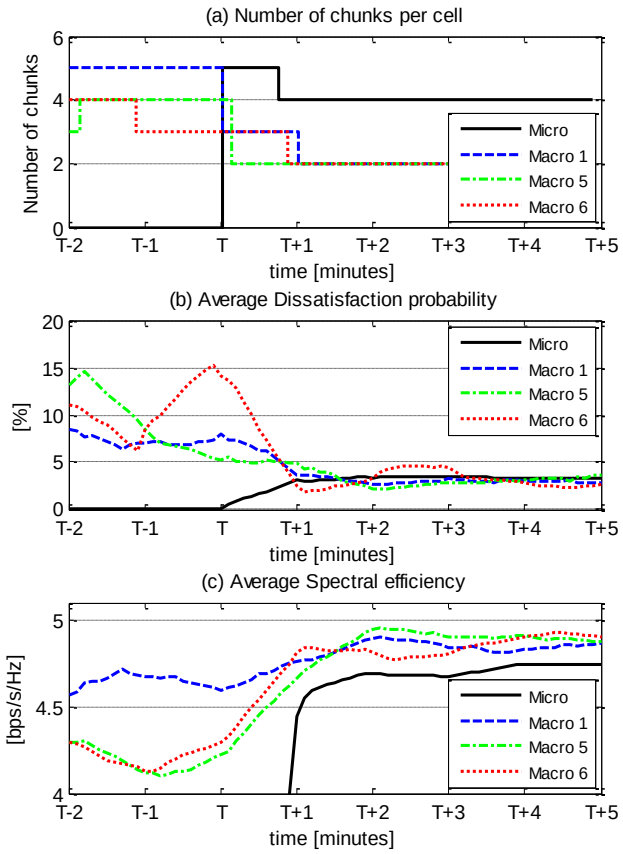


**Figure 7.27 Number of assigned chunks, average dissatisfaction probability and spectral efficiency for the scenario with hot-spot.**

### 7.2.2 Application to a femtocell deployment

In this section the femtocell Scenario C without a macrocell layer presented in Chapter 6 is studied. The performance of the non-cooperative spectrum assignment scheme based on RL-DSA is compared with reference spectrum assignment schemes in femtocell networks given on Chapter 3. Concretely, RL-DSA is compared with Random-FRF1, Random-FRF2, Random-FRF3 and Random-FRF6 strategies (i.e., strategies where the available spectrum is divided into $v$ =1, $v$ =2, $v$ =3 or $v$ =6 portions and femtocells randomly select one of them to operate). As in previous section, the reward signal per cell selected for RL-DSA is based on SINR as given by (5.14) in Chapter 5. RL-DSA configuration values are taken from Table 7.4.

In Scenario C 10 and 100 circular femtocells are randomly deployed with a uniform distribution in a square area of 500x500 m2 to obtain the performance comparison under two different densities of femtocells, and the target user's throughput varied from 128 kbits/s to 2048 kbits/s to examine performance under different traffic loads. Other details can be found in section 6.3.3 in Chapter 6.

Figure 7.28 shows the average performance statistics for the case study in terms of dissatisfaction probability and spectral efficiency for the two densities of femtocells. In both cases, RL-DSA maintains the dissatisfaction probability at the lowest level even for high QoS throughput targets. To this end, RL-

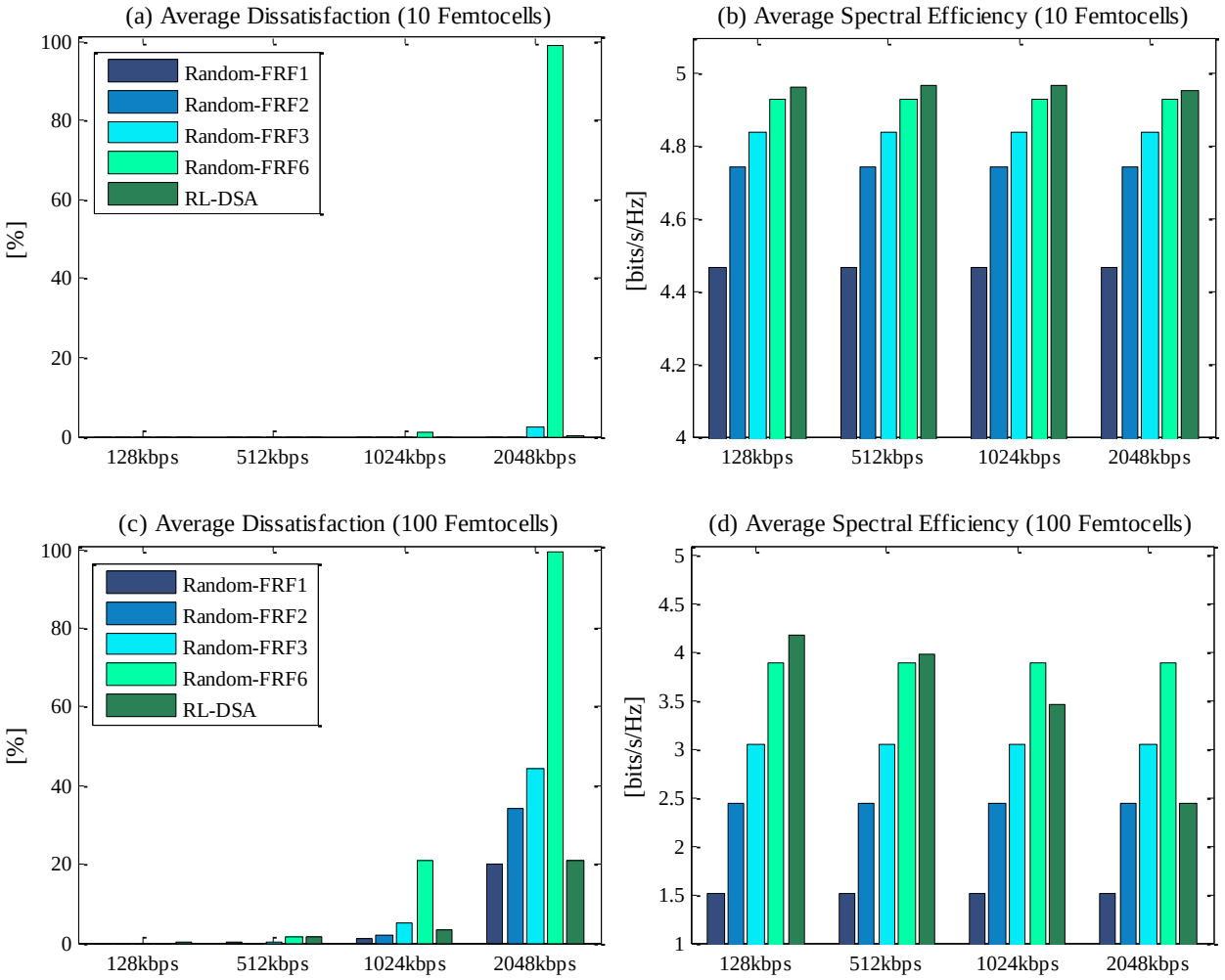


**Figure 7.28 Performance comparison in pure femtocell scenario.**

DSA adapts the number of chunks per cell to cope with the demanded traffic. Logically, an increment in the assigned bandwidth per cell increases the possibility of having intercell interference. Thus, in the high density scenario with strong intercell interference due to cells' proximity (e.g., 100 femtocells) RL-DSA experiments a reduction of the spectral efficiency. On the other hand, the greater the value of $v$ in the Random-FRF, the lower the probability that two adjacent femtocells use the same portion and interfere each other, which turns into an increment of the spectral efficiency. However, a high $v$ reduces the available bandwidth in each femtocell, what leads to an increase of dissatisfaction probability due to the lack of capacity (this is especially remarkable for Random-FRF6 with a target throughput of 2048 kbits/s). In all, the decentralized DSA framework based in RL-DSA is, once more, the strategy that obtains the best tradeoff between dissatisfaction probability and spectral efficiency.

Finally, Figure 7.29 depicts the SINR CDF comparison for 10 and 100 randomly deployed femtocell when users require 512kbps. It can be seen, that certainly, in this case the decentralized DSA scheme ameliorates SINR, what translated into spectral efficiency improvements in Figure 7.28.

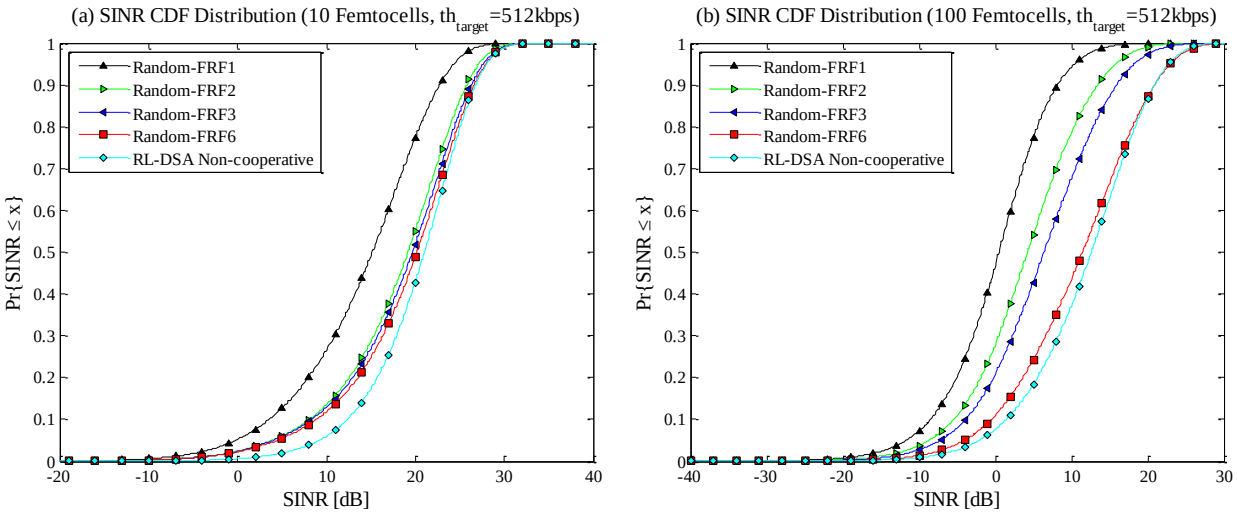


**Figure 7.29 SINR Cumulative Distribution Function comparison in the femtocell scenario**

## 7.2.3 Application to a macrocell and femtocell deployment

This section focuses on a two-layer cellular deployment scenario compound of macrocells and femtocells given by Scenario D in Chapter 6. Co-channel spectrum assignment is considered so that intercell interference between the macrocell and femtocell layer could arise. Performance results are obtained for two deployments of the femtocells.

### 7.2.3.1 Randomly deployed femtocells in a macrocell

The layout of 19 macrocells is combined with 10 or 100 femtocells randomly positioned in the coverage area of the central macrocell. As an example, Figure 7.30 illustrates a sample scenario with 10 femtocells. Performance is obtained for different spectrum assignment strategies in the macrocell and femtocell layer respectively. In particular, FRF1, FRF3 or RL-DSA are tested in macrocell deployment at the same time that Random-FRF1, Random-FRF3 or RL-DSA are used in the femtocell layer, having a total of nine possible combinations. In the case of RL-DSA strategy, the cooperative spectrum assignment

scheme was used at the macrocell layer whereas the non-cooperative scheme was employed at the femtocell layer.

Table 7.7 and Table 7.8 show the average spectral efficiency obtained in the central macrocell and the femtocell layer for the deployment of 10 femtocells and 100 femtocells respectively. The employment of RL-DSA strategies in both the macrocell and femtocell layer brings important spectral efficiency improvements, being such improvement particularly significant at the femtocell layer, and when RL-DSA is used at both the macrocell and femtocell layer. Hence, the inclusion of self-organization at both layers is clearly beneficial.

Finally, Figure 7.31 illustrates the average SINR improvement in the central macrocell when RL-DSA strategy is used. Nine combinations of possible spectrum assignment strategies for both the macrocell and femtocell layer are studied. In this figure, each row represents a spectrum assignment for the macrocell layer whereas each column represents a spectrum assignment for the femtocell layer. Notice that the best average SINR pattern in the cell is obtained when RL-DSA is used at both layers.

Furthermore, femtocells produce SINR *dead zones,* where the SINR for macrocell users highly decays in the proximities of a femtocell. However, certain spectrum assignment schemes, these zones can be avoided. In fact, RL-DSA strategy in both the macrocell and femtocell layer is the best approach, demonstrating a better average SINR pattern in the central macrocell.

**Table 7.7 Spectral efficiency comparison in macrocell-femtocell scenario with 10 femtocells.**

| | | **Macrocell spectral efficiency (bits/s/Hz)** | | | **Femtocell spectral efficiency (bits/s/Hz)** | | |
|---|---|---|---|---|---|---|---|
| | **Macrocell Spectrum assignment** | FRF1 | FRF3 | RL-DSA | FRF1 | FRF3 | RL-DSA |
| **Femtocell Spectrum Assignment** | Random-FRF1 | 3.34 | 5.00 | 5.00 | 2.48 | 3.95 | 4.39 |
| | Random-FRF3 | 3.35 | 5.00 | 5.00 | 2.60 | 4.12 | 4.63 |
| | RL-DSA | 3.35 | 5.00 | 5.00 | 2.59 | 4.64 | 4.95 |

**Table 7.8 Spectral efficiency comparison in macrocell-femtocell scenario with 100 femtocells.**

| | | **Macrocell spectral efficiency (bits/s/Hz)** | | | **Femtocell spectral efficiency (bits/s/Hz)** | | |
|---|---|---|---|---|---|---|---|
| | **Macrocell Spectrum assignment** | FRF1 | FRF3 | RL-DSA | FRF1 | FRF3 | RL-DSA |
| **Femtocell Spectrum Assignment** | Random-FRF1 | 2.94 | 4.28 | 4.28 | 1.56 | 2.19 | 2.42 |
| | Random-FRF3 | 3.15 | 4.85 | 4.85 | 2.10 | 3.23 | 3.73 |
| | RL-DSA | 3.15 | 4.95 | 4.99 | 2.09 | 3.84 | 4.31 |

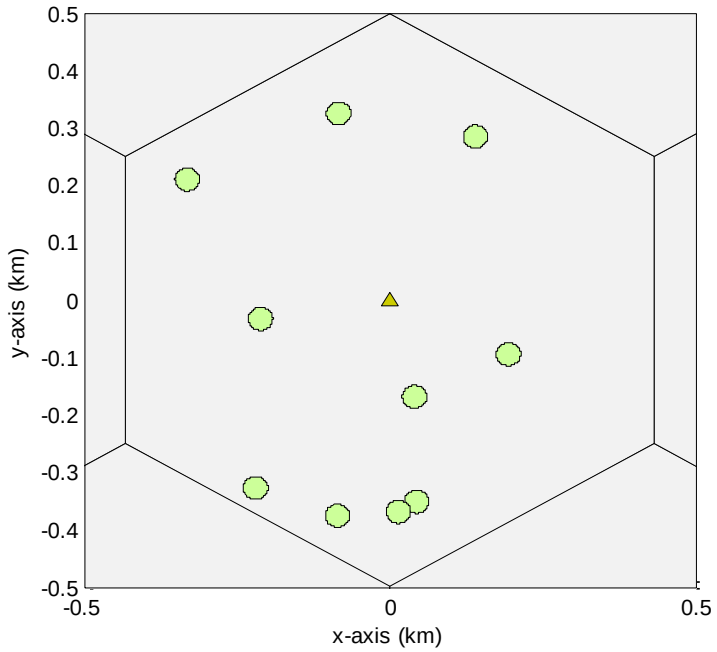


**Figure 7.30 Illustrative picture of the macrocell and femtocell scenario with 10 randomly deployed femtocells.**

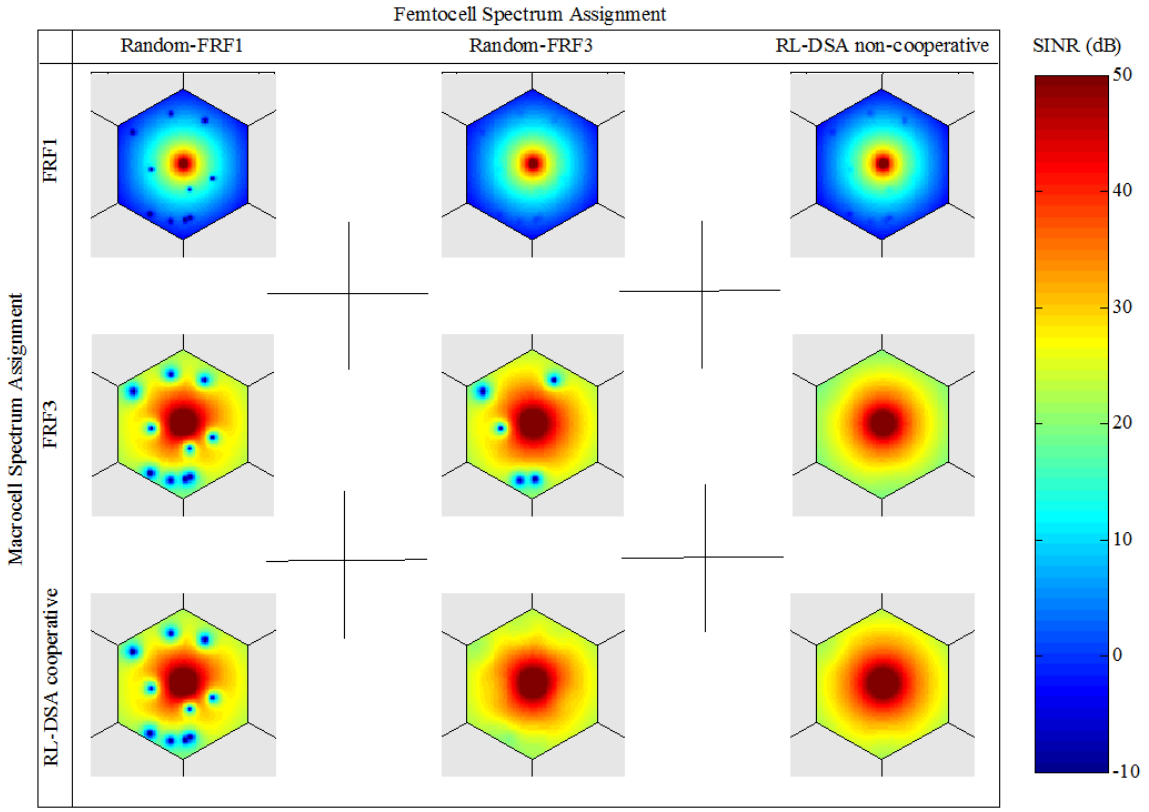


**Figure 7.31 Average SINR comparison in the central macrocell of the sample macrocell-femtocell scenario shown in Figure 7.30.**

#### 7.2.3.2 Indoor deployed femtocells

In this case femtocells are deployed inside a building to test the decentralized DSA framework under a realistic indoor environment. Scenario layout was given in Figure 6.10 of Chapter 6, where focus is done in the central macrocell of the 7-cell cellular layout. That macrocell contains a building with 12 offices where one femtocell is deployed per office to offer indoor coverage.

Two case studies are considered to assess the performance of the algorithms. For comparison purposes, case study 1 considers a femtocell deployment without macrocell layer whereas case study 2 considers also the macrocell deployment with a FRF3 spectrum assignment, i.e., the central macrocell uses one third of the spectrum and the rest of the macrocells alternate one of the other 2 subbands. Basically, this allows evaluating the performance for the femtocell layer under an orthogonal and co-channeling spectrum deployment respectively. For the femtocell layer, the Random-FRF3 scheme is used as reference. As DSA algorithm, in this case the HEUR-DSA1 algorithm is employed, to show that the decentralized framework is also able to self-organize the spectrum arrangement with other DSA strategies different from RL-DSA.

Results are presented (Figure 7.32) in terms of spectral efficiency and user's throughput dissatisfaction probability. Three different thresholds for the satisfaction throughput have been considered such as 512, 1024, 2048 kbits/s, so that the behavior of the Random-FRF3 and HEUR-DSA1 strategies with different QoS requirements are assessed. In the figure solid lines represent case study 1 whereas dashed lines represent case study 2 results.

Figure 7.32(a)(b) shows the performance comparison for both case studies when the homogeneous distribution of the users in the building is employed. There, both case studies exhibit the same qualitative performance, but, case study 2 attains lower spectral efficiency and slightly higher the dissatisfaction probability. This is because of the presence of the macrocell layer, which increases the intercell interference. Comparing the Random-FRF3 and HEUR-DSA1 strategies, the HEUR-DSA1 scheme demonstrates the best trade-off between QoS fulfillment and spectral efficiency. For instance, for 512 kbits/s satisfaction throughput, the HEUR-DSA1 scheme obtains the best spectral efficiency with a reduced dissatisfaction probability. On the other hand, for 2048 kbits/s satisfaction throughput, each femtocell demands a higher number of subchannels to cope with the traffic demand, what translates into a higher intercell interference and hence into a reduction of the spectral efficiency. However, compared with the random strategy, the HEUR-DSA1 scheme reduces considerably the dissatisfaction probability.

Besides, it is important to highlight the effect of a heterogeneous spatial distribution of the traffic load as shown in Figure 7.32(c)(d). There, the benefits of the HEUR-DSA1 scheme are appreciable even with lower satisfaction throughputs than in the homogeneous case. Note that, in general, a heterogeneous distribution of the load will be common in real scenarios. Thus, this calls for using adaptive approaches such as the proposed decentralized DSA framework.

As an example of the SINR improvements that self-organization could bring to femtocell deployments, Figure 7.33 shows the SINR distribution in the proximities of the building for both case studies. The spectrum assignments for the Random-FRF3 and HEUR-DSA1 schemes in the homogeneous distribution with 512 kbits/s are considered (analogous results have been found for the other tests). Points outside the building are taken as if they were connected to the central macrocell when applicable (case study 2).

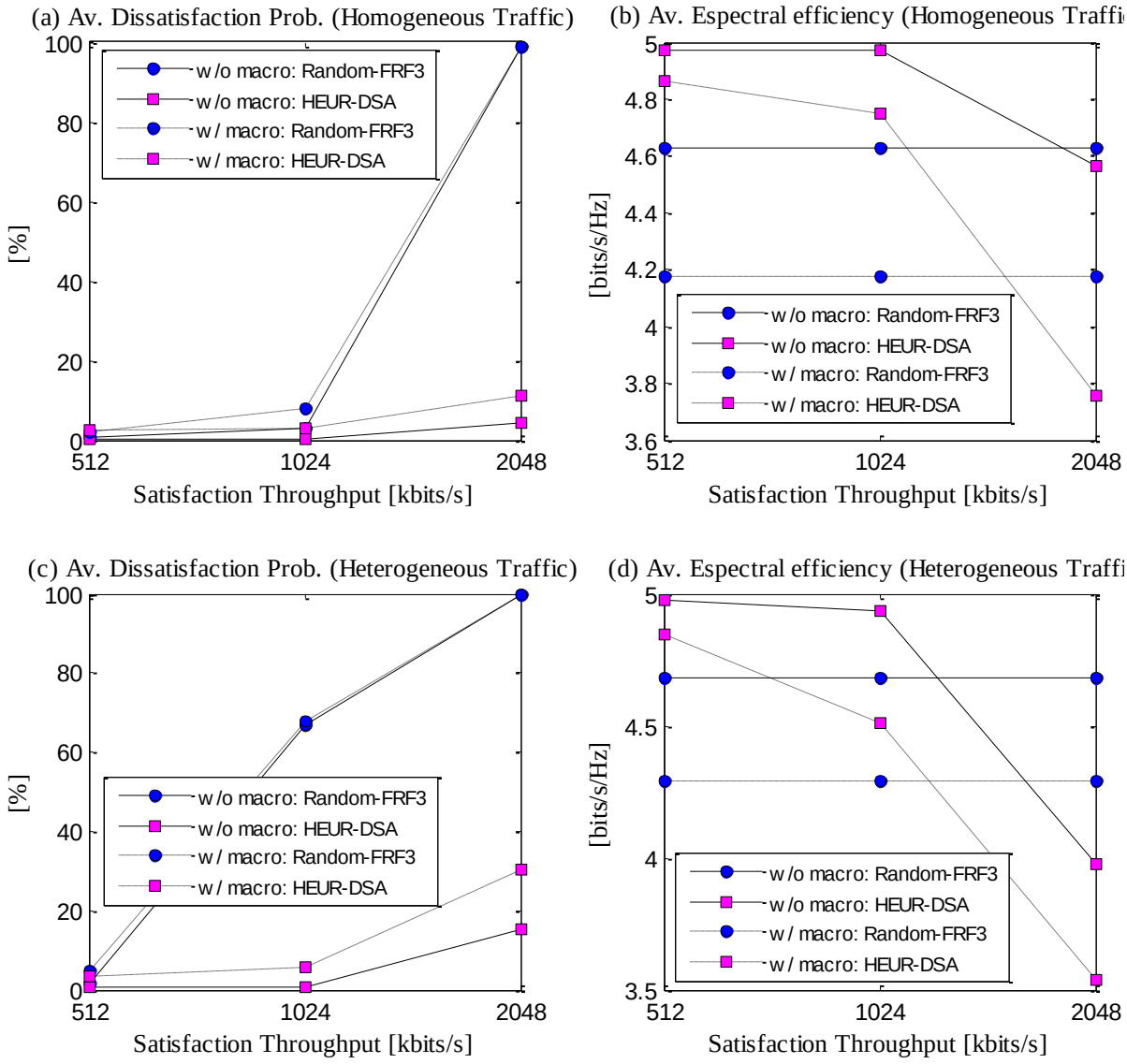


**Figure 7.32 Performance comparison for combined macrocell and femtocell scenario with indoor femtocell deployment.**

It can be seen in Figure 7.33 that, the self-organized scheme considerably ameliorates overall femtocells SINR with respect to the Random-FRF3 scheme. Concretely, in case study 2 the self-organized scheme does not create interference in the macrocell layer (brown color) whereas the Random-FRF3 scheme reduces in several dBs the SINR in the building's surroundings. It has been determined that in the former, femtocell layer self-organizes so that each femtocell uses different subchannels to those used by the central macrocell. However, it is appreciable a considerable reduction of the SINR in the femtocells compared with case study 1 (i.e., when macrocell layer was not present). It has been checked that, for each femtocell, this is mainly because of the interference produced by the other macrocells, rather than by other femtocells. Hence, in general, the interference from macrocells distinct of the macrocell where femtocells are deployed cannot be neglected.

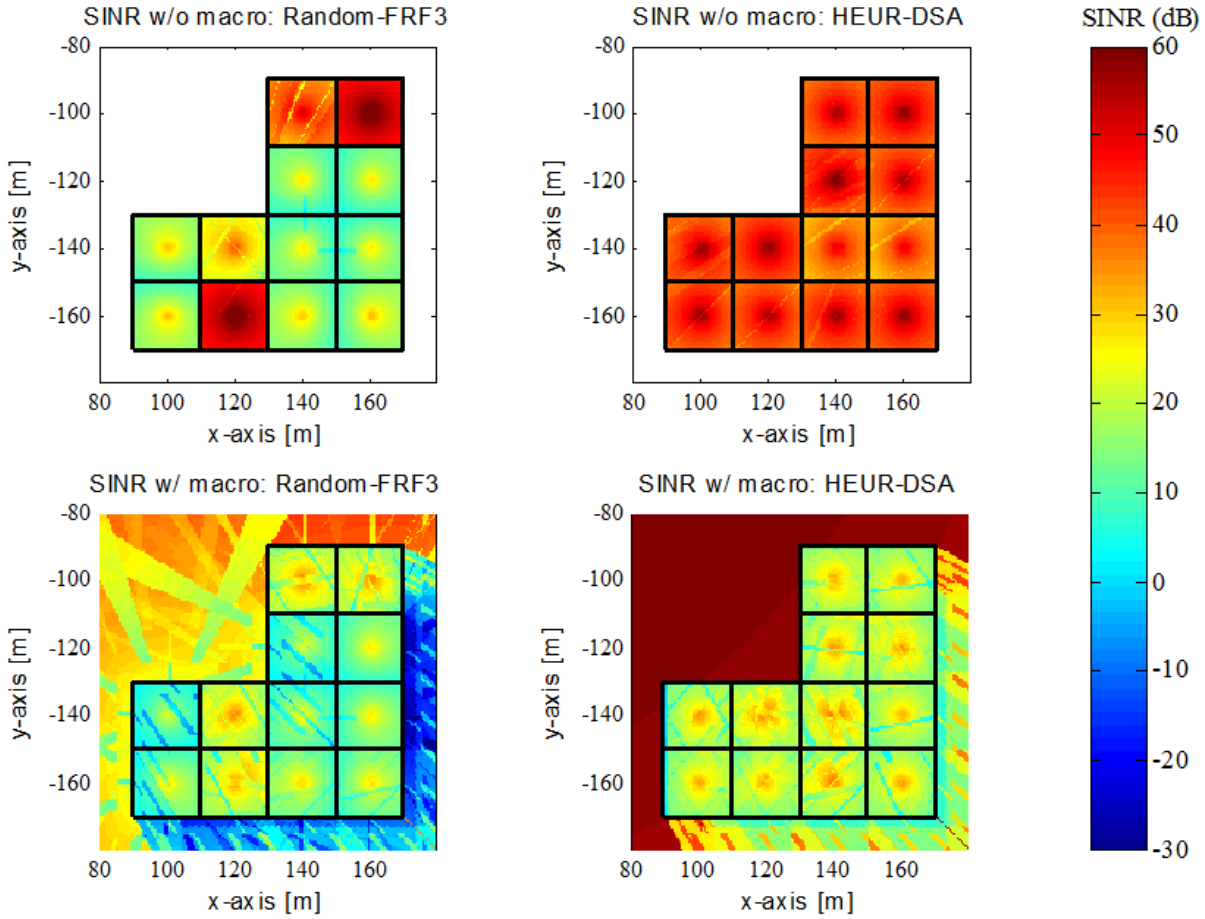


**Figure 7.33 SINR comparison in the surroundings of the building**

Finally, Figure 7.34 illustrates the dynamics of the decentralized DSA self-organized framework for case study 1 with homogeneous traffic and 512 kbits/s. The figure shows how the femtocells in the building self-organize the spectrum so that the spectral efficiency is improved. The figure shows the average spectral efficiency evolution for 7 time instants after femtocells switch-on. Notice that, after 18 executions of the algorithms, the spectral efficiency is near maximum for all cells. In this case, the value of the execution period is $L$=10s, so that the self-organization takes around 3 minutes, which is an appropriate time considering that, usually, traffic demands vary in a similar time-scale.

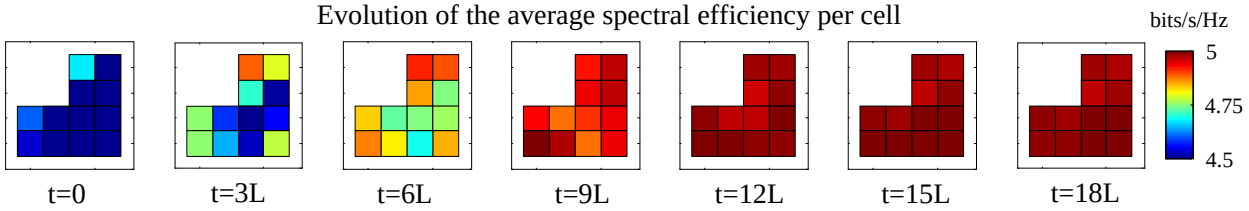


**Figure 7.34 Dynamics of the DSA framework. Self-organization of spectrum to improve the average spectral efficiency per femtocell.**

## 7.3 Summary

This chapter has presented an exhaustive study of the performance attained by the proposed DSA framework and DSA algorithms in Chapter 4 and Chapter 5 respectively for OFDMA cellular networks. Performance evaluation has been carried out in the four scenarios presented in Chapter 6, which encompass two macrocell deployments, a decentralized femtocell deployment and a combined macrocell and femtocell deployment respectively. In those scenarios, results were obtained by performance comparison with state-of-the-art strategies such as Frequency Reuse Factor (FRF), Partial frequency Reuse (PR), and Soft frequency Reuse (SR) for macrocell deployments and Random-FRF as reference spectrum assignment strategy in femtocell deployments.

The centralized version of the DSA framework has been first tested over a macrocell scenario with different static spatial distributions of the traffic load. There, the four variants of the HEUR-DSA algorithms have been shown their capability to adapt the spectrum assignment to the different spatial traffic patterns. These algorithms ameliorate overall system's spectral efficiency at least 33% respect to the total reuse of frequency resources in the cellular system. At the same time, the HEUR-DSA algorithms maintain or improve the dissatisfaction probability up to 20% for users at the edge of the cell. It has also been demonstrated that the proposed HEUR-DSA algorithms enable the releasing of spectrum bands in large geographical areas so that this spectrum can be exploited by secondary users. This property has been assessed under several spatial load distributions including the homogeneous case, although the best performance is obtained under highly heterogeneous spatial distributions. Furthermore, it has been shown that the greater the freedom that the DSA algorithms have to perform the chunk-to-cell assignment, the better the performance obtained. Thus, HEUR-DSA1 and especially HEUR-DSA2, which benefit from more flexibility in the chunk assignment, show better performance than HEUR-DSA3 and HEUR-DSA4, which include constraints in the chunks to be assigned to edge or central users. In any case, all the four considered strategies overcome the performance of classical PR and SR strategies.

Following with the same simulation scenario with static spatial distribution of the traffic load, the performance of the RL-DSA algorithm has been evaluated, revealing a superior performance in spectral efficiency and dissatisfaction probability to HEUR-DSA, certainly because of being an optimal algorithm. Also, RL-DSA has exhibited an improvement of up to 10 MHz*km2 in Useful Released Surface over HEUR-DSA. Illustrative OFDMA chunks usage maps over the macrocell scenario indicate that RL-DSA is able to release a given chunk in a vest area of contiguous cells, being then suitable for secondary usage without causing harmful interference to primary communications.

Then, the dynamism and adaptability of the centralized DSA framework in a macrocell scenario where the traffic load is dynamically varied has been examined Performance comparison results for FRF1, FRF3, PR, SR, HEUR-DSA2 and RL-DSA confirm that dynamic strategies achieve the best tradeoff between spectral efficiency and user's throughput dissatisfaction probability. Also, throughput fairness is attained, achieving such a dynamic strategies a better balance between the throughputs for users in the scenario than the obtained with fixed reuse factor strategies. It has been also illustrated how the DSA framework reacts to a traffic load variation in the scenario. Concretely, for an increment of the traffic load, an increment of the dissatisfaction probability is reported. Then, the DSA framework executes the DSA algorithm to search for a new spectrum assignment and then drastically reduce the dissatisfaction probability.

The second part of the chapter has presented the performance evaluation for the decentralized DSA framework in the context of downlink OFDMA cellular networks. The proposed framework has shown its effectiveness in several cellular deployments such as a macrocell deployment, a femtocell deployment and a combined macrocell and femtocell deployment. Results in terms of spectral efficiency and SINR improvements and QoS fulfillment confirm the added value of DSA strategies and concretely, the suitability of the decentralized DSA approach. For instance, it has been shown that the non-cooperative spectrum assignment scheme is quite suitable for decentralized and user-deployed femtocell networks where intercell signaling interfaces are not feasible. However, this kind of interface is being considered for an operator-deployed macrocell network. In this scenario, the cooperative spectrum assignment scheme has shown better performance than the non-cooperative scheme.

Apart from performance improvements, several scenarios have shown the potential benefits of implementing a decentralized self-organized DSA framework. For example, it has been shown how a microcell could automatically self-configure its initial spectrum assignment after switch-on to adapt to the spectrum usage in neighboring macrocells and to avoid intercell interference. This has been also tested in a femtocell scenario, where femtocells self-organize the spectrum usage to increase spectral efficiency after several periodic executions of the DSA algorithm. Moreover, in two-layer deployments it has been determined that the effect of neighboring of the macrocell where femtocells are deployed cannot be neglected. That is, those macrocells could negatively interfere femtocells producing a performance reduction. Hence, multicell macrocell deployments should be used in two-layer macrocell and femtocell performance evaluations, rather than single-cell macrocell plus femtocells, as usual in certain studies.

Finally, important SINR improvements have been revealed in two-layer deployments, where the usage of the decentralized framework, in both the macrocell and femtocell layer, can help to avoid dead coverage zones in the macrocell layer around femtocells due to strong cross-layer interference. Then, the macrocell and femtocell layers self-organize to select different spectrum assignments to avoid such cross-layer interference. The interesting point is that no explicit cooperation between those layers is permitted and thus, the proposed decentralized framework is certainly a practical solution to solve the deployment of thousands of femtocells in an already deployed macrocell environment.

# 8

# CONCLUSIONS

## Outline

# 8 Conclusions

Spectrum assignment in the context of multicell OFDMA 4G mobile networks is a very challenging task. Apart from the usual objective of intercell interference mitigation to obtain the highest spectral efficiency per OFDMA frequency resource, other challenges arise such as facing (*i*) the high decentralization degree of the network due to the appearance of new femtocell deployments, (*ii*) the temporal and spatial variations of the traffic load, or (*iii*) the new spectrum regulatory models, all them opening new research lines in the field.

This thesis has contributed to this research by proposing a solution to dynamically manage the cell-by-cell spectrum assignment in an OFDMA cellular deployment. Particularly, a full Dynamic Spectrum Assignment (DSA) solution has been proposed opposed to classical fixed frequency planning where OFDMA frequency resources are statically assigned to cells.

Therefore, our proposal mainly has contributed to:

1.- Efficiently manage intercell interference, by assigning frequency resources to cells in way that Signal to Interference plus Noise Ratio per OFDMA frequency resource is increased, thus considerably improving the attainable spectral efficiency in terms of bits per second and per unit of bandwidth.

2.- Cope with variable temporal and spatial traffic load in medium-term, so that the frequency resources given to each cell are adapted to users' QoS requirements.

3.- Exploit future regulatory spectrum frameworks where primary (licensees) and secondary users of the spectrum co-exist in the same geographical area. Then, pieces of unnecessary primary spectrum in a given region at a given point of the time are pooled and released to be used by a secondary spectrum market, thus increasing the overall spectrum usage.

4.- Ease the spectrum assignment in future decentralized femtocells scenarios, where a high degree of independency on the decision taken by nodes regarding radio resource management and, in particular, the usage of OFDMA frequency resources is needed.

To this end, two major proposals have been contributed during the work on this thesis.

On the one hand, a practical framework to implement DSA in a next generation mobile cellular network has been defined. The following major features characterize the framework:

- The operation of the DSA framework has been divided into two time-scales, so that in the medium-term the cell-by-cell spectrum assignment is decided, whereas in the short-term frequency OFDMA resources are assigned to users. This brings two major advantages. On the one hand, the framework is prepared to either adapt to variations of the traffic load in the medium-term, or to fast channel variations in the short-term to exploit frequency and multiuser diversity gains. On the other hand, the decoupling between the temporal scales allows reducing the complexity of the DSA strategies, which operate in a virtually unconstrained environment in terms of time employed to take decisions. That is, the medium-term execution of the DSA strategies relaxes the signaling and computational requirements of the DSA strategies.

- A centralized and decentralized functional architecture can be configured for the framework. This generalizes the framework so that it can be applied to a vast number of scenarios, from typical macrocell scenarios where centralized control is employed, to future femtocell scenarios where nodes are almost independent. The centralized scenario benefits of handling global information at the time of a new spectrum assignment decision. Besides, it has been proved that, thanks to the medium-term execution, signaling requirements between the centralized execution entity and the cells are completely assumable for current mobile networks infrastructure. In the case of the decentralized version of the DSA framework, a cooperative and a non-cooperative spectrum assignment scheme were proposed where inputs to the framework come from an intercell signaling interface or from measurements taken in the own cell respectively. In this latter case, full independent execution of the framework at a cell level is achieved.

- The framework is open to implement several DSA strategies, since a modular architecture has been defined, where a specific functional block for a DSA algorithm has been included. Also, the rest of supporting entities and interfaces have been defined. In particular, a supporting network/cell characterization entity has been proposed where an offline interaction between it and the DSA algorithm has been designed. This is because this brings two major benefits when facing a spectrum assignment optimization task. First, the physical time taken by the real network/cell to return a proper averaged response for a spectrum assignment can be unacceptably slow in terms of elapsed time compared with the quick response that network/cell characterization entity could provide. Second, candidate solution spectrum assignments can, in some cases, suppose a prohibitive cost in performance for the live network if they are applied on-line. On the contrary, applying those candidate spectrum assignments to an off-line entity only suppose a simulation cost, so, virtually, any candidate spectrum assignments can be tested.

- Last but not least, self-organization has been taken as the building stone of the framework, so that either the network (in the case of the centralized approach) or each particular cell (in the case of the decentralized approach) can autonomously react to changes in the spectrum requirements. That is, the framework is able to observe the network/cell environment, analyze its status, and automatically trigger the appropriate decision mechanisms to re-adapt the spectrum assignment. Certainly, compared with current human-driven frequency planning tasks, this approach constitutes a practical contribution to reduce operational expenditures in a mobile 4G network.

On the other hand, this thesis has proposed two types of DSA algorithms to address the problem spectrum assignment in OFDMA-based radio interfaces and to show the flexibility of the proposed framework, which is able to adopt, in its envisaged functional architecture, different implementations of the DSA algorithm.

The first type of algorithm, named HEUR-DSA, is based on heuristics methodology. It is a simple but effective algorithm that was introduced to intuitively solve the problem of assigning frequency resources to a set of cells depending on the traffic load needs per cell. Then, the HEUR-DSA algorithm first computes the number of chunks needed per cell and then selects the most suitable ones according to intercell interference costs. Four flavors of the algorithm have been presented, where some of them

incorporate self-tuning mechanisms to better adapt the spectrum usage, whereas others are designed to improve the throughput fairness between the users in the cell.

The second type of algorithm, named RL-DSA, is based on reinforcement learning methodology. The inclusion of learning procedures to manage the spectrum in an OFDMA radio interface definitely constitutes another major contribution of this work. Also, the specific REINFORCE learning methodology is a novel point of view in the wireless field. It is a low complexity method that demonstrates an appealing optimization behavior of a given reward signal. Concretely, we have exploited this inherent feature of RL to propose RL-DSA, which, after a number of steps selects a spectrum assignment that maximizes a given reward signal conveniently defined in terms of target performance metrics. Then, different rewards signals addressing SINR or spectral efficiency optimization have been defined. Finally, studies about convergence behavior show an excellent robustness of RL-DSA with respect to the values selected for its parameters.

It is worth to remark the low complexity of the proposed DSA algorithms. Only few simple multiplications and additions are needed for both HEUR-DSA and RL-DSA. Also the usage of REINFORCE for RL-DSA avoids the need of huge action-value databases as in other RL methodologies such as Q-learning. Instead, REINFORCE simplifies memory requirements by using few records to store internal weights and recent values taken by the reward signal.

Also, notice that another property of the DSA framework is its flexibility to change the optimization objective of the framework. For instance, one only needs to change the reward signal and the corresponding network/cell characterization functional block devoted to build such a reward signal to change the optimization target maintaining the same DSA algorithm.

Performance results to assess the validity of proposed DSA solutions have been carried out in several scenarios comprising two macrocell deployments, a decentralized femtocell deployment and a combined macrocell and femtocell deployment respectively. In those scenarios, results were obtained by performance comparison with state-of-the-art strategies for macrocell deployments and femtocell deployments. Obtained results reveal the following benefits of implementing proposed DSA solutions:

- Overall system's spectral efficiency is improved at least 33% respect to the total reuse of frequency resources in the cellular system in some scenarios. At the same time, user's throughput dissatisfaction probability, taken as a QoS measure, is maintained or improved up to 20% for users at the edge of the cell in some cases.

- Throughput fairness is attained, where DSA strategies achieve a better balance between the throughputs for users in the scenario than the obtained with fixed reuse factor strategies.

- Proposed DSA framework and algorithms are able to release spectrum bands in large geographical areas so that this spectrum can be exploited by secondary users. In particular, DSA strategies has exhibited an improvement of up to 15 MHz*km2 in Useful Released Surface over fixed strategies over certain scenarios. This means that DSA is able to release a given chunk in a vast area of contiguous cells, being then suitable for secondary usage without causing harmful interference to primary communications.

- Results in terms of spectral efficiency and SINR improvements, and QoS fulfillment confirm the added value and the suitability of the decentralized DSA approach for decentralized and user-deployed femtocell networks, where intercell signaling interfaces are not feasible.

- In two layer deployments composed of macrocells and femtocells, the usage of the DSA strategies at both layers allows them to select spectrum assignments to avoid cross-layer interference, without explicit cooperation between layers.

These properties have been obtained under several spatial load distributions including the homogeneous case, although the best performance is obtained under heterogeneous spatial distributions, since fixed spectrum assignment schemes are not adapted for those scenarios. On the other hand, a heterogeneous distribution of the traffic load is the most common situation in a mobile environment.

In all, this thesis claims that the proposed DSA framework can be very promising for future 4G wireless networks based on OFDMA technology since: (1) primary operator's spectral efficiency is improved, without degrading primary users' QoS, (2) new opportunities for secondary markets are created, which opens new business chances for spectrum holders and operators to obtain a profitable return from their investments in the expensive spectrum licenses. Then, this framework can be a feasible solution to gradually evolve current networks and their spectrum regulatory rules in the context of e.g. the private commons paradigm, (3) it incorporates self-organization concepts that bring important performance improvements and operational costs reductions, not only from the point of view of spectrum assignment optimization but also from the perspective of initial deployment, where nodes are able to self-configure after switch-on introducing a minimal impact on already deployed system, and, hence, (4) it is certainly a practical solution to solve the deployment of thousands of femtocells in an already deployed macrocell environment.

## 8.1 Future work

The work of this thesis has focused on dynamic spectrum assignment strategies over a downlink radio interface based on OFDMA, where a full-buffer traffic service has been employed to analyze the performance of different spectrum assignment strategies mainly in terms of achievable capacity or throughput (i.e., satisfaction of minimum user's throughput requirements and also spectral sufficiency or throughput per unit of bandwidth).

Then, a first line of extension of the thesis' work could be to study the performance of current implementation of the DSA framework with other traffic services such as real-time, streaming, or delay-constrained services. Also, the impact of other short -term schedulers different from proportional fair or round robin could be examined. Thus, the implementation of network/cell characterization entity could require to be revised to accommodate those traffic services and schedulers. However, the functional architecture of the DSA algorithms and the framework could remain unaltered, showing in that case the flexibility of the proposed mechanisms in this thesis.

Furthermore, the utility of the DSA framework and algorithms for operation on uplink direction could be evaluated. Roughly, it seems that the proposed framework could be applied in uplink, since input metrics needed by Status Observer can be directly estimated by the base station even without the need of measurement reporting form users. Then, the framework operation could be followed as for downlink. However, some modifications regarding restrictions on the way that OFDMA frequency resources have to be assigned to cells have to be probably considered, such us the need that the assigned chunks are contiguous to reduce the Peak-to-Average-Power Ratio in terminals.

In a more general point of view, future work could focus on future trends in mobile communications. For instance, future wireless networks tend to be heterogeneous, where the user has a key role making use of heterogeneous devices in heterogeneous environments with heterogeneous access networks. As a result, from the point of view of a mobile network's operator, it faces the management of a network that comprises several radio access technologies (e.g., 3GPP GERAN, UTRAN and future LTE). Each one of these radio access technologies (RAT) requires their own radio resource management (RRM) strategies [81] to optimize their behavior (e.g., handovers, power control, etc.). Moreover, it has been studied that certain RRM functionalities can be jointly managed across the radio access technologies, constituting the so called Joint Radio Resource Management (JRRM) strategies [82]. These strategies aim manage functions such as initial RAT selection (i.e., the functionality devoted to decide to which RAT a given service request should be allocated at the set-up phase) and the vertical handover (i.e., the functionality devoted to decide a seamless RAT switching for an on-going service).

Given this context, another possible future research line to expand the work of this thesis could be to extend the framework to cover the spectrum assignment over different RATs. Then, DSA framework could take into account JRRM operation, since JRRM decisions certainly affect the traffic load per RAT and cell, affecting as well to the spectrum requirements in the network. On the other hand, if lack of spectrum is detected for certain cell of a given RAT, JRRM could be instantiated by DSA framework to handoff part of the load in that RAT in that specific cell to another RAT with available spectrum.

Finally, under current regulatory spectrum framework, each RAT is deployed in a given spectrum band that is exclusive for the services offered for that RAT. However, and as already pointed out at the beginning of this dissertation, the regulatory perspective in how the spectrum should be managed is evolving to a more flexible mindset, where spectrum is used without regulatory and technological barriers. For instance, *spectrum refarming* is seen as an initial step of such a flexible scenario. Refarming attends to the possibility of exploiting the most convenient RAT on the most convenient spectrum band within the spectrum bands licensed to an operator.

This opens a new horizon for future work, where the DSA framework is not only limited to manage the spectrum of a single RAT in a multi-RAT network, but the spectrum of multiple RAT simultaneously, considering the all the spectrum bands licensed to the operator as a common pool where the cell-by-cell spectrum assignment could be done in a more flexible manner. For sure, aspects such as the guard bands between spectrum and the different size of frequency channels used by different RATs have to be taken into account.

# Bibliography


[1] Recommendation ITU-R M.1645, Framework and overall objectives of the future development of IMT-2000 and systems beyond IMT-2000, 2003

[2] Resolution ITU-R 57, "Principles for the process of development of IMT-Advanced", 2007

[3] V. Chandrasekhar, J. Andrews, A. Gatherer, "Femtocell networks: a survey," IEEE Commun. Mag., vol.46, no.9, pp.59-67, September 2008

[4] ABI Research, "Femtocell Access Points: Fixed-Mobile Convergence for Residential, SMB, and Enterprise Markets", Aug. 2006

[5] E. Bogenfeld et al. "Collaborative Cognitive RRM, Dynamic Spectrum Management, and Self-Organisation for Future Mobile Networks", ICT Mobile and Wireless Communications Summit, Florence, Italy, June 2010.

[6] M.M. Buddhikot, "Understanding Dynamic Spectrum Access: Models, Taxonomy and Challenges," 2nd IEEE International Symposium on in New Frontiers in Dynamic Spectrum Access Networks (DySPAN) 2007, pp. 649-663. DOI: 10.1109/DYSPAN.2007.88

[7] S. Haykin, "Cognitive Radio: Brain Empowered Wireless Communications" *IEEE Journal on Selected Areas in Communications*, Vol. 23, No. 2, February 2005.

[8] IBM Research Headquarters (manifesto). Autonomic computing: IBM's perspective on the state of information technology. [Online]. Available: http://www.research.ibm.com/autonomic/overview/elements.html. October 2001.

[9] J. Jang, K.B. Lee, "Transmit power adaptation for multiuser OFDM systems," IEEE J. Sel. Areas Commun., 2003, 21(2), pp. 171- 178

[10] F. Khan, LTE for 4G Mobile Broadband, Cambridge University Press, 2009.

[11] R. van Nee, R. Prasad, "OFDM for Wireless Multimedia Communications", Artech House, 2000. ISBN:0890065306

[12] S. Pietrzyk, "OFDMA for Broadband Wireless Access", Artech House, 2006. ISBN:9781596930445

[13] 3GPP TS 45.005 v8.5.0, "Technical Specification Group GSM/EDGE Radio Access Network; Radio transmission and reception" Mayo 2009.

[14] 3GPP TS 36.300 v8.11.0 "Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2" Jan 2010.

[15] IEEE 802.16 Working Group on Broadband Wireless Access, http://WirelessMAN.org/

[16] IEEE. Standard 802.16-2004. Part16: Air interface for fixed broadband wireless access systems. October 2004.

[17] IEEE. Standard 802.16e-2005. Part16: Air interface for fixed and mobile broadband wireless access systems—Amendment for physical and medium access control layers for combined fixed and mobile operation in licensed band. December 2005

[18] C. Prehofer, C. Bettstetter, "Self-organization in communication networks: principles and design paradigms," IEEE Commun. Mag., vol.43, no.7, pp. 78-85, July 2005

[19] A.G. Spilling, A.R. Nix, M.A. Beach, T.J. Harrold, "Self-organisation in future mobile communications," Electronics & Communication Engineering Journal , vol.12, no.3, pp.133-147, Jun 2000

[20] S. Dixit, E. Yanmaz, O.K Tonguz, "On the design of self-organized cellular wireless networks," IEEE Comm.Mag., vol.43, no.7, pp. 86-93, July 2005

[21] M. Amirijoo, et al., "Use Cases, Requirements and Assessment Criteria for Future Self-Organising Radio Access Networks" Lecture Notes in Computer Science, Ed. Springer Berlin / Heidelberg, pp 275-280, 2008

[22] E. Bogenfeld, I. Gaspard (editors) "Self-x in Radio Access Networks," Available at: https://www.ict-e3.eu/project/dissemination/whitepapers/whitepapers.html Dec. 2008.

[23] INFSO-ICT-216284 Self-Optimisation and Self-Configuration in Wireless Networks (SOCRATES) Project, http://www.fp7-socrates.org/?q=node/1

[24] ICT-2007-216248 End-to-End Efficiency (E3) Project, http://www.ict-e3.eu/

[25] R. S. Sutton, A. G. Barto, "Reinforcement Learning: An Introduction", A Bradford Book, The MIT Press, 1998, ISBN 0-262-19398-1

[26] L. P. Kaelbling, M. L. Littman, and A. P. Moore, "Reinforcement learning: A survey," Journal of Artificial Intelligence Research, vol. 4, pp. 237-285, 1996.

[27] C. J. C. H. Watkins and P. Dayan, "Technical note: Q-learning," Machine Learning, vol. V8, no. 3, pp. 279-292, May 1992.

[28] R. J. Williams, "Simple statistical gradient-following algorithms for connectionist reinforcement learning" Machine Learning, Springer Netherlands. 8(3-4), May 1992

[29] V. V. Phansalkar, and M. A. Thathachar, "Local and global optimization algorithms for generalized learning automata". Neural Computation. Vol. 7(5) pp. 950-973, Sep. 1995

[30] M.A.L Thathachar, P.S. Sastry, "Varieties of learning automata: an overview", IEEE Trans. on Man, and Cybernetics, Part B: Cybernetics, Vol 32(6) pp. 711- 722, Dec 2002

[31] S. Chieochan, and E. Hossain, "Adaptive radio resource allocation in OFDMA systems: a survey of the state-of-the-art approaches", Wireless Communications & Mobile Computing, Volume 9 , Issue 4 (April 2009), Pages: 513-527

[32] S. Pietrzyk, G.J.M Janssen. "Subcarrier allocation and power control for QoS provision in the presence of CCI for the downlink of cellular OFDMA systems" VTC-Spring, 22-25 April 2003, pp. 2221- 2225 vol.4

[33] G. Kulkarni, S. Adlakha, M. Srivastava,"Subcarrier Allocation and Bit Loading Algorithms for OFDMA-Based Wireless Networks",IEEE Transactions on Mobile Computing, Nov. 2005, 4(6), pp. 652-662

[34] K. Kim, H. Kim, Y. Han, "Subcarrier and Power Allocation in OFDMA Systems", 60th IEEE Vehicular Technology Conference-Fall pp. 1058-1062[Note(s) : 7 vol., ] 2004

[35] L. Yan, Z. Wenan, S. Junde, "An Adaptive Subcarrier, Bit and Power Allocation Algorithm for multiuser ODFM systems", CCECE 2003

[36] J. Heo, I. Cha, K.H. Chang, "Effective adaptive transmit power allocation algorithm considering dynamic channel allocation in reuse partitioning-based OFDMA system", Wireless Pers Commun 2007, 43, pp 677-684

[37] L. Jorguseski, R. Prasad, "Downlink Resource Allocation in Beyond 3G OFDMA cellular systems", PIMRC 2007

[38] G. Li, H. Liu, "Downlink Radio Resource Allocation for Multi-Cell OFDMA System", Trans Wireless Comm, 2006, 5(12), 3451-3459

[39] H. Kwon, W.-I. Lee, and B. G. Lee , "Low-Overhead Resource Allocation with Load Balancing in Multi-cell OFDMA Systems", Vehicular Technology Conference, 2005. VTC 2005-Spring. 2005 IEEE 61st Volume 5, 30 May-1 June 2005 Page(s):3063 – 3067

[40] Z. Wang; R.A, Stirling-Gallacher,., "Frequency reuse scheme for cellular OFDM systems," Electronics Letters , vol.38, no.8, pp.387-388, 11 Apr 2002

[41] M. Mustonen, K.Hooli, J.Ylitalo, A. Tölli, "Application of Intra-RAN Flexible Spectrum Use to Networks with Multi-Rate Services", 17th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'06), 2006.

[42] M. Sternad, T.Ottosson, A. Ahlen, A. Svensson, "Attaining both Coverage and High Spectral Efficiency with Adaptive OFDM Downlinks", IEEE Vehicular Technology Conference VTC2003-Fall, Orlando, Florida, Oct. 2003.

[43] S-E. Elayoubi, B. Fourestie, "On frequency allocation in 3G LTE systems", 17th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'06), 2006.

[44] Huawei, "Soft Frequency Reuse Scheme for UTRAN LTE", 3GPP R1-050507 May 2005

[45] H. Kim, Y. Han, J. Koo, “Optimal Subchannel Allocation Scheme in Multicell OFDMA Systems”, in IEEE 59th Vehicular Technology Conference vol. 3, 2004, pp. 1821-1825.

[46] Y. J. Choi, C. S. Kim, and S. Bahk, "Flexible Design of Frequency Reuse Factor in OFDMA Cellular Networks," in ICC'06: IEEE International Conference on Communications. vol. 4, 2006, pp. 1784-1788.

[47] C. Lengoumbi, P. Godlewski, P. Martins "Dynamic subcarrier Reuse with Rate Guaranty in a Downlink Multicell OFDMA system" in PIMC06.

[48] S-E. Elayoubi, O. Ben Haddada, B. Fourestie, “Performance Evaluation of frequency Planning schemes in OFDMA-based Networks”, IEEE Transactions on Wireless Communications, May 2008; 7(5) pp. 1623-1633.

[49] Y. Xiang, J. Luo, C. Hartmann, “Inter-cell Interference Mitigation through Flexible Resource Reuse in OFDMA based Comunication Networks”, EW2007

[50] D. López-Pérez, A. Jüttner, and J. Zhang, “Dynamic Frequency Planning Versus Frequency Re-use Schemes in OFDMA Networks” IEEE Vehicular Technology Conference, Barcelona, Spain, 26-29 Apr., 2009.

[51] D. López-Pérez, A. Valcarce, G. de la Roche; J. Zhang, “OFDMA femtocells: a roadmap on interference avoidance,” IEEE Comm. Mag., vol.47, no.9, pp.41-48, Sept. 2009

[52] J. Espino, J. Markendahl, “Analysis of Macro - Femtocell Interference and Implications for Spectrum Allocation”, 20th Personal, Indoor and Mobile Radio Communications Symposium 2009 (PIMRC-09) Tokyo, Japan, 13-16 September, 2009

[53] J. Góra, T. E. Kolding, “Deployment Aspects of 3G Femtocells”, 20th Personal, Indoor and Mobile Radio Communications Symposium 2009 (PIMRC-09) Tokyo, Japan, 13-16 September, 2009

[54] Y. Bai, J. Zhou, L. Liu, L. Chen, H. Otsuka, “Resource Coordination and Interference Mitigation between Macrocell and Femtocell”, 20th Personal, Indoor and Mobile Radio Communications Symposium 2009 (PIMRC-09) Tokyo, Japan, 13-16 September, 2009

[55] Y.-Y. Li, M. Macuha, E. S. Sousa, T. Sato and M. Nanri, “Cognitive Interference Management in 3G Femtocells” 20th Personal, Indoor and Mobile Radio Communications Symposium 2009 (PIMRC-09) Tokyo, Japan, 13-16 September, 2009

[56] H. Claussen, D. Calin, “Macrocell Offloading Benefits in Joint Macro and Femtocell Deployments”, 20th Personal, Indoor and Mobile Radio Communications Symposium 2009 (PIMRC-09) Tokyo, Japan, 13-16 September, 2009

[57] V. Chandrasekhar, J.G. Andrews, “Spectrum allocation in tiered cellular networks”,. IEEE Trans. on Comm., vol. 57(10) pp. 3059 - 3068 Oct. 2009

[58] H. Claussen, L. T. W. Ho, L. G. Samuel, “Self-optimization of coverage for femtocell deployments,” in Proc. Wireless Telecommunications Symposium (WTS), 2008.

[59] Y.-S. Chen, C.-J. Chang, F.-C. Ren, “Q-learning-based multirate transmission control scheme for RRM in multimedia WCDMA systems,” IEEE Trans. on Veh. Tech., vol. 53(1), pp. 38-48, Jan. 2004,

[60] L. Giupponi, R. Agustí, J. Pérez-Romero, O.Sallent, “Fuzzy Neural Control for Economic-Driven Radio Resource Management in Beyond 3G Networks”, IEEE Trans. on Systems, Man, and Cybernetics—Part C: Applications and Reviews, vol. 39(2), pp.170-189, March 2009

[61] H. Tong, T. X. Brown, “Reinforcement Learning for Call Admission Control and Routing under Quality of Service Constraints in Multimedia Networks”, Machine Learning, vol. 49(2-3) pp. 111 – 139, Nov-Dec 2002

[62] T. Chee-Kin Hui, C.-K. Tham, “Adaptive Provisioning of Differentiated Services Networks Based on Reinforcement Learning”, IEEE Trans. on Systems, Man, and Cybernetics—Part C: Applications and Reviews, vol. 33(4), pp.492-501, Nov. 2003

[63] J. Nie and S. Haykin, "A dynamic channel assignment policy through q-learning", IEEE Transactions on Neural Networks, vol. 10, no. 6, pp. 1443-1455, 1999.

[64] J. Nie and S. Haykin, "A q-learning-based dynamic channel assignment technique for mobile communication systems," IEEE Transactions on Vehicular Technology,vol. 48, no. 5, pp. 1676-1687, 1999.

[65] N. Lilith and K. Dogancay, "Dynamic channel allocation for mobile cellular traffic using reduced-state reinforcement learning," in Wireless Communications and Networking Conference, 2004. WCNC. 2004 IEEE, vol. 4, 2004, pp. 2195-2200

[66] S. Singh and D. Bertsekas, "Reinforcement learning for dynamic channel allocation in cellular telephone systems," in Advances in Neural Information Processing Systems, vol. 9, The MIT Press, 1997.

[67] S. M. Senouci and G. Pujolle, "Dynamic channel assignment in cellular networks: a reinforcement learning solution," in 10th International Conference on Telecommunications, 2003. ICT 2003., vol. 1, 2003, pp. 302-309

[68] A. J. Goldsmith, S. Chua, "Variable-rate variable-power MQAM for fading channels" IEEE Trans. Commun., vol. 45, pp. 1218—1230, 1997

[69] R. Schoenen, R. Halfmann, B.H. Walke, "MAC Performance of a 3GPP-LTE Multihop Cellular Network" IEEE International Conference on Communications, 2008. ICC '08, pp.4819-4824, May 2008

[70] C. Wengerter, J. Ohlhorst, A.G.E. von Elbwart, "Fairness and throughput analysis for generalized proportional fair frequency scheduling in OFDMA," IEEE 61st VTC 2005-Spring. Jun. 2005.

[71] 3GPP TS 45.005 v8.5.0, "Technical Specification Group GSM/EDGE Radio Access Network; Radio transmission and reception" Mayo 2009.

[72] A.L. Stolyar, H. Viswanathan, "Self-Organizing Dynamic Fractional Frequency Reuse in OFDMA Systems," IEEE Conference on Computer Communications. INFOCOM 2008. pp.691-699, 13-18 April 2008

[73] S. Xinghua, H. Zhiqiang, N. Kai, W. Weiling, "A Hierarchical Resource Allocation for OFDMA Distributed Wireless Communication Systems," IEEE Global Telecommunications Conference, 2007. GLOBECOM '07, pp.5195-5199, 26-30 Nov. 2007

[74] J. Choi, S. Bahk, "Cell-Throughput Analysis of the Proportional Fair Scheduler in the Single-Cell Environment", IEEE Trans. on Vehicular Technology, vol. 56, pp. 766-778, 2007

[75] L. Berger et al. Interaction of transmit diversity and proportional fair scheduling, 57th IEEE Vehicular Technology Conference, 2003. VTC 2003-Spring., 2003, vol.4, 2423-2427.

[76] ETSI TR 101 112, "UMTS: Selection procedures for the choice of radio transmission technologies of the UMTS", Version 3.2.0, April 1998.

[77] COST (European Co-operation in the Field of Scientific and Technical Research), COST 231 Book, Final Report. Chapter 4, Propagation Prediction Models.

[78] 3GPP TSG-RAN1#48, "LTE physical layer framework performance verification", R1-070674, February 2007

[79] R. Irmer (Ed), "Next Generation Mobile Networks Radio Access Performance Evaluation Methodology", NGMN white paper, June 2007

[80] J. Nasreddine, J. Pérez-Romero, O. Sallent, R. Agustí, "A primary spectrum management solution facilitating secondary usage exploitation". 17th ICT mobile and wireless communications summit June 2008.

[81] J. Pérez-Romero, O. Sallent, R. Agustí, M. A. Díaz-Guerra, Radio resource management strategies in UMTS, John Wiley & Sons, 2005

[82] L. Giupponi, "Contribution on Joint Radio Resource Management in Beyond 3G Networks", PhD Doctoral Dissertation. Universitat Politècnica de Catalunya, May 2007