The Islamic University - Gaza
Research & Graduate Affairs
Faculty of Engineering
Master of Electrical Engineering Department

Interference Management in OFDMA Femtocell Network

Submitted by:
Salim Ahmed Mahdi

Supervised by:

Dr. Fady El-Nahal **Dr. Musbah Shaat**

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master in Electrical Engineering-Communication Systems.

2014

TABLE OF CONTENTS

Abstract

The next generation cellular wireless network aims to efficiently deploy low cost and low power cellular base station in the subscriber's indoor environment. Also, one of the effective techniques of improving the coverage and enhancing the capacity and data rate in cellular networks is to reduce the cell size and transmission distances. Therefore, the concept of deploying femtocells over macrocell has recently attracted growing interests in academia and in telecommunication operators.

Interference management becomes a major issue in the deployment of femtocells because they share the same licensed frequency spectrum with macrocell, so there are interference between neighboring femtocells called Co-tier interference and between the femtocell and macrocell called Cross-tier interference. Wherefore, the demand increases for mitigation interference techniques. Femtocell may not be possible to carry out an elaborate frequency planning of the femtocell network so they are expected to have self-configuring ability.

In this study, we proposed an adaptive Co-tier and Cross-tier interference avoidance scheme by self-organization using reinforcement learning in orthogonal frequency division multiple access (OFDMA). We developed an autonomous radio resource (RB) allocation algorithm to pursue the most efficient frequency allocation. In order that, self-organized power and resources block allocation technique solved the interference problem caused by a femtocell network operating in the same channel in cellular network.

We consider modeling the femto network as a multi agent system where each femto base stations is the agent in charge of managing the radio resources to be allocated to their femtousers. So, real-time multi-agent reinforcement learning, known as decentralized Q-learning, to manage the interference generate to femto/ macro-users with no more traffic in backbone network. In order that, the agent is interacting with the surrounding environment in a distributed fashion so that is able to learn an optimal policy to solve the interference problem.

الملخص

يهدف الجيل المستقبلي للشبكة اللاسلكية الخلوية نشر محطات خلوية بكفاءة عالية و منخفضة التكلفة و ذات استهلاك طاقة أقل. ومن الطرق الفعالة لتحسين و زيادة كفاءة الخدمة للمشترك، من نقل بيانات واتصالات وخدمات أخرى، هي تقليل مساحة تغطية المحطة الخلوية و تقليل المسافة بين المستخدم و المحطة و زيادة عدد المحطات. ولذلك، فإن مفهوم الفيمتوسيل و هي الوحدات المصغرة من المحطات الخلوية التي يتم نشرها و تنصيبها ضمن نطاق الماكروسيل ، المحطات الخلوية ذات تغطية اكبر ، قد جذبت مؤخرا اهتمام متزايد من الأوساط الأكاديمية ومشغلي الاتصالات السلكية واللاسلكية.

بسبب اشتراك المحطات الصغيرة الفيمتوسيل بنفس الطيف الترددي للمحطات الكبيرة الماكروسيل تنبثق صعوبات منها التداخل بين المحطات و من هنا تصبح إدارة التداخلات القضية الرئيسية في توزيع الفيمتوسيل ضمن نطاق. و هناك نوعين اثنين من التداخل أولهما ، تداخل بين محطات الفيمتوسيل المتجاورة و يطلق عليها تداخل متشاركة الطبقة ، و النوع الآخر تسمى تداخل تقاطع الطبقة و هي بين محطة الفيمتوسيل و محطة الماكروسيل. ولهذا السبب، تزايدت الطلبات على تقنيات تخفيق و ادارة التداخلات. ليس هناك أية إمكانية لتنفيذ تخطيط لتوزيع الأطياف الترددية للفيمتوسيل لطبيعة تركيب و توزيع محطة الفيمتوسيل، و لذلك من المتوقع تواجد تقنيات للإدارة الذاتية لتهيئة الموارد من طيف ترددي و توزيع للطاقة.

في هذه الدراسة، فإننا سوف نقترح مشروع لتقليل التداخل بنوعيهما المذكورين آنفاً عن طريق التنظيم الذاتي باستخدام تعزيز التعلم في نظام تقسيم الترددات المتعامدة للوصول المتعدد. و طورنا خورازمية لتوزيع الطيف الترددي بشكل مستقل و بدون أية تخطيط مسبق و بكفاءة عالية. من أجل ذلك، الإدارة الذاتية للموارد حلت مشكلة التداخل الذي تسببه شبكات الفيمتوسيل التي تعمل ضمن الشبكات الأخرى.

وفي هذه الاطروحة نفترض نموذج لشبكة فيمتوسيل كنظام متعدد الوكلاء حيث كل محطة فيمتوسيل هي الوكيل المسؤول عن إدارة الموارد الراديوية التي ستخصص لمستخدمي شبكة الفيمتوسيل. لذلك، هذا النظام يحتاج وقت لتعزيز التعلم لعدة وكلاء ، والمعروفة باسم التعلم اللامركزي، و لحل مشكلة التداخلات دون زيادة الحمل على شبكة النظام الكلي. من أجل ذلك، فإن كل وكيل يتفاعل مع البيئة المحيطة بطريقة توزيعها بحيث يكون قادرا على معرفة السياسة المثلى لحل مشكلة التداخل.

All praise goes to Allah, the Creator and Lord of the Universe

and the Cause of Every Success in My Whole life

This Work is dedicated to …

My Great Parents,

My Dear Wife, and

Lovely Brothers and Kind Sister.

\

ACKNOWLEDGEMENTS

In the name of Allah the most Compassionate and the most Merciful who has favored me with countless blessings.

It is really a great moment to reach this part of the thesis. When you are writing the "acknowledgement" means that you are done, and you can take a deep breath and look back through the days passed while you are stepping a stair up to remember and thank all people who gave a significant push and made this step possible.

In fact, the major thanks and gratitude are owned to my supervisor, Dr. Musbah Shaat, his advices, kind directions, guides and scientific assistant that draw and cleared the right way to reach this achievement were, always, the backbone alike.

I would like also to offer my heartfelt thanks to my co-supervisor, Dr. Fadi Alnahal, who believed on me, provided me with support and always welcomed my questions with the best possible guidance and hints.

A lot of thanks are to the Islamic University of Gaza library which facilitates the accessibility to different important global libraries that made the research simpler and efficient.

Great thanks to my parents for their support and prayers that blessed my life. Loving thanks to my patient and kind wife Dr.Hanan who lived the pressure with me in the past two years, my brother Islam who was the best friend and gave me his heartfelt advises and support, my warm-hearted and lovely sister Samer.

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

1G	First Generation
2G	Second Generation
3G	third Generation
3GPP	third Generation Partnership Project
4G	Fourth Generation
ACI	Adjacent Channel Interference
ASE	Area Spectral Efficiency
ADC	Analog-to-Digital Converter
BS	Base Station
CAPEX	CAPital EXpenditure
CCI	Co-Channel Interference
CDMA	Code Division Multiple Access
CP	Cyclic Prefix
CR	Cognitive Radio
CI	Computational Intelligence
CoI	Cell of Interest
CSG	Closed Subscriber Group
DAC	Digital- to-Analogue Converter
DCA	Dynamic Channel Allocation
DFS	Dynamic Frequency Selection
DSL	Digital Subscriber Line
eNB	eNode B
EPC	Evolved Packet Core
EPS	Evolved Packet System
FBS	Femtocell Base Station

FCC	Federal Communications Commission
FDD	Frequency Division Duplex
FFT	Fast Fourier Transformation
HSS	Home Subscriber Service
IEEE	Institute of Electrical and Electronics Engineers
ISI	Inter-Symbol Interference
IFFT	Inverse Fast Fourier Transformation
LTE	Long Term Evolution
MA	Multiple Access
MDP	Markov Decision Processes
MIMO	Multiple Input Multiple Output
MME	Mobility Management Entity
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency-Division Multiple Access
PAPR	Peak-to-Average Power Ratio
PCRF	Policy and Charging Rules Function
PDN	Packet Data Network
P-GW	Packet Data Network Gateway
PU	Primary User
QoS	Quality of Service
RAN	Radio Access Network
RB	Resource Block
SAE	System Architecture Evolution
SC	Sub Carrier
S-GW	Serving Gateway
SNR	Signal to Noise Ratio

SU	Secondary User
TDD	Time Division Duplex
TTI	Transmission Time Interval
UE	User Equipment
WLAN	Wireless Local Access Network

Chapter One

Introduction

1.1 Introduction

The growth of mobile data traffic in the upcoming years is expected. The number mobile broadband subscribers are foreseen to reach 3 billion worldwide by 2016 shown Figure-1.1 [1]. Yet, indoor users suffer the most because the transmissions originate from base stations (BS) positioned outdoors and must pass through walls before reaching their destination.

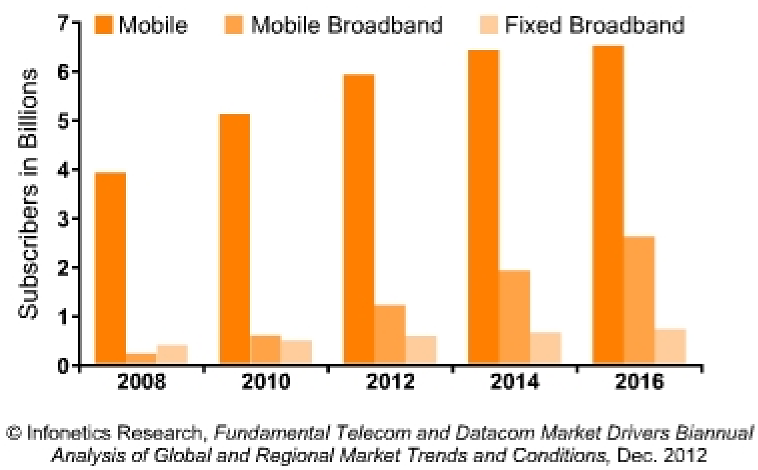


Figure 1.1: Subscribers Trends [1].

Simply deploying more macrocells is infeasible to alleviate the increasing traffic load. Through reduced cell sizes and transmit distance is how the highest increase in wireless capacity is achieved. This gain comes out from higher spectral efficiency. In addition, allocating cells inside buildings will enhance the QoS for indoor users.

Femtocells, home evolved Node-B (HeNB), are short-range low-cost BS installed by the consumer to enhance the wireless coverage indoors. The user-installed femtocell device communicates with the cellular network over a broadband connection such as Digital Subscriber Line (DSL). Recently, the femtocells are introduced as one of the candidates by the Third Generation Partnership Project (3GPP) Long-Term Evolution (LTE) [2].

Since the femtocells are designed to be installed by customers in a plug and play manner, the location and number of the femtocells are unknown to the network operator. Subsequently, femtocells may be randomly distributed depending on customer's requirements, so there may exist a large number of neighboring femtocells in densely populated areas; this scenario may cause a radio coverage of femtocell overlapping with other radio coverage of neighboring femtocells.

The co-channel deployment is adopted to enhance the resource efficiencies, where the same frequency-time resources are occupied by different femtocells transmitters so that the unwanted signal may be sent from the neighboring femtocells which decreases the communication quality, known as the co-tire interference problem .In addition, the employ femtocell underlay macro cell's coverage causes interference on the macro user called cross-tire interference [3]. The co-tire and cross-tier interference cause significant degradations of the network performance which have been widely studied in the literature

OFDMA femtocells have several advantages over code-division multiple access (CDMA). OFDMA systems divided spectrum into orthogonal subcarriers which provide more ability for the avoidance of the intracell interference in addition to the robustness to multipath. OFDMA works as a multi-access technique by allocating different users to different groups of orthogonal subchannels. Moreover, OFDMA femtocells can exploit channel variations in both frequency and time domains for the avoidance of interference [4].

Self-organization was first presented by Ashby in [5] and lately defined as the global order developing from local interactions, where the benefit of self-organization for the system is evolving behaviors of node without beforehand planning actions and changing its structure and function as a result of the sum of all the interactions of its components and the environment as a whole. Currently, self-organizing capabilities intended in wireless systems that can be classified in self-configuration, self-optimization and self-healing.

Reinforcement learning "Q-learning" is learning what to do, how to map situations to actions, so as to maximize a numerical reward signal. Furthermore, reinforcement learning is defined not by characterizing learning methods, but by characterizing a learning problem. The learner or the agent is not told which actions to take, as in most forms of machine learning, but the optimal actions must discovered by trying them. In the most interesting and challenging cases, actions may affect not only the immediate reward but also the next situation and, through that, all subsequent rewards. These two characteristics trial-and error search and delayed reward are the two most important distinguishing features of reinforcement learning [6].

This thesis agreements with the introduction of femtocells in heterogeneous networks where heterogeneous network mean the users or wireless nodes distributed in non-uniform spatial distribution. In more detail, some of the above mentioned challenges are considered and combined, resulting in an intercell interference coordination approach based on self-organization techniques, as detailed in the following sections.

1.2 Motivation

In the recent years, new technologies, like smarts phone, tablets and smart home etc., have grown rapidly. Thus, telecommunications companies are interested in low-cost technologies and work efficiently to cover the increased users. From here emerged the idea of home base station (femtocell) where 2/3 of calls and over 90% of data services occur indoors. So, Femtocells are expected to substantially reduce the operator CAPital EXpenditure (CAPEX) and Operational EXpenditure (OPEX) [13]. Therefore, new challenges face this industry because the femtocell worked as plug and play. Subsequently, this scenario causes a radio coverage of femtocell overlapping with other radio coverage of neighboring femtocells and underlying with macrocell. As a result, the interference will decrease the efficiency of this technology where femtocells share same license frequency of macrocells.

Self-organization techniques based on Machine Learning (ML) approaches are introduced to perform the self-management (configuration, optimization and self-healing) for femtocell that supports implementation of femtocell in procedures for coexistence of macro and femto networks. In particular, we propose to map the femtocells onto a multiagent system [6], where each femto BS is an intelligent and autonomous agent that learns [14] by directly interacting with the environment and by properly utilizing the past experience. The environment in which the multiagent system is operating is dynamic due to the characteristics of the mobile wireless scenario.

Inspired by all of the above benefits, in this thesis, we propose an intelligent machine learning technique based on reinforcement learning to self-management of power and frequency resources for femtocell base station configuration where the constrains are mitigating overall interferences and enhancing the QoS.

1.3 Objectives

The main objectives for this study can be summarized in the following points:

1. To study the main theoretical background required for the research topic .
2. Analyzing the femtocell system in general and focusing on the radio resources and interference management in particular.
3. Studying the details and challenges related to the radio resources management and its impact on the mitigation of the interference and efficiency of femtocell networks.
4. Investigating the impact learning techniques for the different modelling scenarios.
5. Propose an efficient algorithm in order to mitigate co-tier and cross-tier interferences.

1.4 Literature Review

The femtocells are introduced as one of the candidates by the Third Generation Partnership Project (3GPP) Long-Term Evolution (LTE) [7]. Serving as the small-scale base stations, the femtocells are adopted to improve the network throughput by extending the coverage into domestic areas such as residences, apartments, offices, and hotspots so the dead zones can be covered and the spectral utilizations can be enhanced for cellular systems. Instead of deploying more macrocells, the deployment of femtocells is an economical option due to its low cost and low power consumption [8]. Since the femtocells are designed to be purchased and deployed by customers, the number and locations of the femtocells are generally unknown to the network operator. Numerous femtocells may be randomly distributed [8]. But considering that there are interferences and handovers among the multiple femtocell base stations, the enterprise femtocell needs a different deployment solution, which configures and optimizes the femtocell accounting the effects of surrounding femtocells and avoiding much interference to them, jointly providing the desired coverage and capacity. Intercell interference cancellation techniques to mitigate the effects of intercell and co-channel interference was suggested in [8]. Moreover, a dynamic selection of predefined antenna patterns has been introduced to reduce the outdoor power leakages as in [9]. However, the above hardware-based approaches usually require the complex and expense of increased hardware cost. On the other hand, power control algorithms and radio resources management provide cost effective approaches in cellular systems for mitigating interference. The interference problem can be solved with adjusting the transmit power but the communication quality might be degraded due to the power reduction. From the perspective of the resource management, an effective approach to allocate different resources with different user requirements directly is presented in [9]. Other researchers in OFDM radio resource management suggested the insertion of sensing frames to scan the whole wireless resources in the current channel [10]. Thus, the femtocells allocate and adjust the usage of the wireless resources by the sensed results. However, the receiving and transmitting data will be interrupted because of the frequent sensing operation. 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 [15]. In the field of mobile cellular networks, several tasks have been identified to adjust network parameters including self-configuration in pre-operational state, self-optimization in operational state, and self-healing in case of failure of a network element [15]. That is, each entity performs its operation based only on the information retrieved from other entities in its vicinity. Hence, self-organization clearly takes a relevant role in the context of femtocells networks. For instance, researchers in [4] propose a self-optimization scheme for frequency planning in the context of OFDMA femtocells. It has been agreed that OFDMA radio access interfaces offer appealing properties such as robustness against multipath fading and high spectral flexibility. Then, as shown in [6], OFDMA FCs facilitate the development of such dynamic self-organization mechanisms, and proof of that is that they are being included in the latest specifications for the LTE system.

We will propose an adaptive co-tire interference avoidance scheme of self-organization in OFDMA femtocell networks by using machine learning scheme (Q-learning).

1.5 Thesis Structure

This dissertation is organized in four chapters; the first one, the introduction, gives a smooth entrance to the topics studied throughout the rest chapters. Architecture, characteristics, advantages, drawbacks, and the current main applications of this technique. In Chapter 2, a brief background of the LTE based in OFDM technique and the machine learning are presented. Chapter 4 describes the system model, problem formulation, Q-learning algorithms that perform the resource allocation, and finally the simulation results and discussion. Chapter 4 comes to a conclusion that summarizes the important issues drawn out from this study and the recommendations on future work to be carried on this subject.

Bibliography

[1] "Telecom-and-Datacom-Market-Drivers-Highlights" [Online]. Available: http://www.infonetics.com/pr/2012/Telecom-and-Datacom-Market-Drivers-Highlights.asp, 2014.

[2] "Picocells-and-Femtocells-to-Be-Part-of-Initial-3G-LTE-Architecture" [Online]. Available: http://www.3gpp.org/news-events/12-news-events-others/press-clippings/1138-Picocells-and-Femtocells-to-Be-Part-of-Initial-3G-LTE-Architecture ,2014.

[3] N. Saquib, E. Hossain, L. Baole and D. I. Kim,"Interference Management of OFDMA in Femtocell Networks: Issues and Approach," IEEE Wireless Communications. June 2012.

[4] Lopez-Perez, D. Valcarce, A. de la Roche and G. Jie Zhang," OFDMA femtocells: a roadmap on interference avoidance ".IEEE Communications Magazine, Volume: 47. September 2009.

[5] W. R. Ashby, "Principles of the self-organizing dynamic system," The Journal of General Psychology, vol. 37, no. 2, pp. 125–128, 1947.

[6] Richard S. Sutton and Andrew G. Barto ,"Reinforcement Learning: An Introduction" ,The MIT Press. Cambridge, Massachusetts London, England, 2005.

[7] Service requirements for Home Node B (HNB) ,3GPP Technical Specification TS 22.220 V11.0.0, Jan. 2011. [Online]. Available: http://www.3gpp.org , 2014.

[8] V. Chandrasekhar, J. G. Andrews, and A. Gatherer, "Femtocell networks: a survey ", Communications Magazine, IEEE Volume: 46, Issue: 9, 2008.

[9]" LTE Overview " [Online]. Available: http:// www.3gpp.org/technologies/keyword sacronyms/98-lte, 2014.

[10] Hamza, J., Long Term Evolution (LTE) - A Tutorial, Network System Labora- tory, Simon Fraser University, Canada, October 13, 2009

[11] A. Z. Yonis," Downlink and Uplink Physical Channels in Long Term Evolution", I.J. Information Technology and Computer Science, 2012, 11, 1-10

[12] Abd El-Gawad , M. Tantawy and El-Mahallawy," LTE QOS Dynamic Resource Block Allocation with power source limitation and queue stability constrains " International Journal of Computer Networks & Communications (IJCNC) Vol.5, No.3, May 2013.

[13] Jie Zhang, Guillaume de la Roche," Femtocells: Technologies and Deployment ", 2010 John Wiley & Sons Ltd, ISBN: 978-0-470-74298-3

[14] Eduardo F. Morales, ”Introduction Reinforcement Learning”, National Institute of Astrophysics, Optics and Electronics, México,2012.

[15] C. Prehofer and C. Bettstetter, “Self-organization in communication networks: principles and design paradigms,” IEEE Commun. Mag., vol. 43, no. 7, pp. 78–85, Jul. 2005.

Chapter two

Background

2.1 OFDM TECHNIQUE

2.1.1 Downlink – OFDMA Overview

In OFDMA, orthogonal narrowband subcarriers can be shared between multiple users. The OFDMA uses sub-carriers and they transmit at low power, unlike full transmission for the whole frequency band. The OFDMA provides good system performance with desired high data rates. In this scheme, the spectrum is basically divided into a series of uniform orthogonal narrowband sub-carriers; with each sub-carrier spaced at 15 KHz with frame duration of $t_f = 10$ ms and it is divided into equal size sub-frame, called Transmission Time Interval (TTI), lasting 1 ms. In addition, the whole bandwidth is divided into 180 kHz physical RBs, each time slot lasting 0.5 ms consisting of 6 or 7 symbols (according to the OFDM prefix-code duration) , therefore one subframe has 1 ms time duration consists two time slot as shown in figure 2.1 [1].

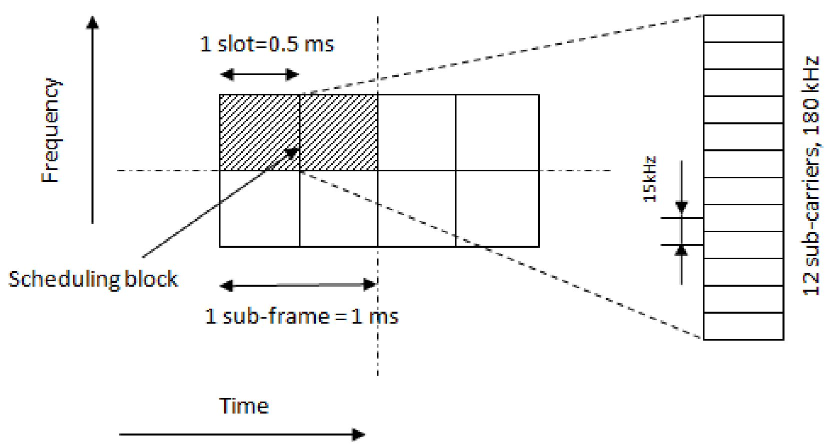


Figure 2.1: Illustration a scheduling of Resource block in OFDMA structure [1].

Transmitted through a time dispersive channel raises two difficulties. First, the system may sometimes transmit multiple OFDM symbols in a series causing intersymbol interference (ISI) between successive OFDM symbols. Second, channel dispersion destroys the orthogonality between subcarriers and cause inter carrier interference (ICI) for the signal. In order that, the peak of one sub-carrier of the spectrum must coincide with the nulls of the other sub-carriers, this is illustrated in figure 2.2 [2]. Furthermore, Inter Symbol Interference problem and the orthogonality problems solved by copy of the last part of OFDM symbol which is appended to the front the transmitted OFDM symbol which is called Cyclic Prefix (CP) [3]. There are two type of CPs, the first one is the Normal CP, which has 7 symbols per slot and the other one is extended CP with 6 symbols per slot. The extended CP has fewer symbols than normal CP because in LTE, CP-length under the Normal CP is 5.21 µs (symbol 0) with short CP of length 4.69 µs ,for symbols 1-6, and for the length of extended CP is 16.7 µs (which is same for all symbols 0-5). To be able to combine multipath components, the CP duration should be longer than the delay spread of the channel as illustrate in figure 2.3 [4].

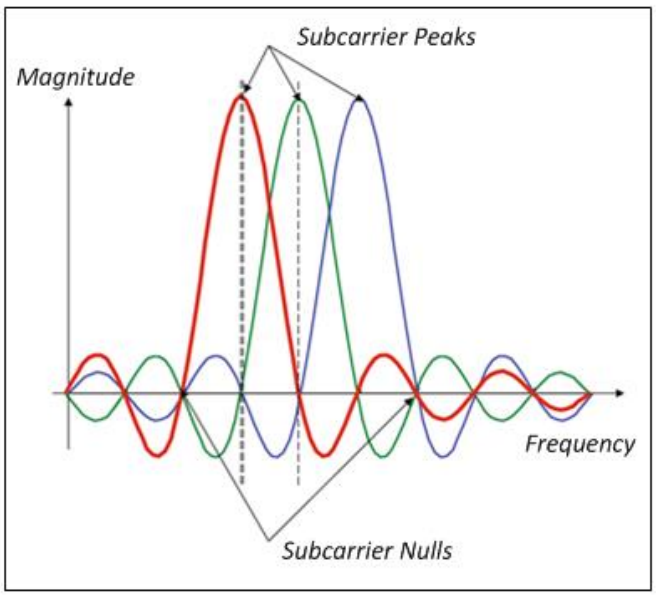


Figure 2.2: Illustrate of Subcarriers in OFDM systems [2].

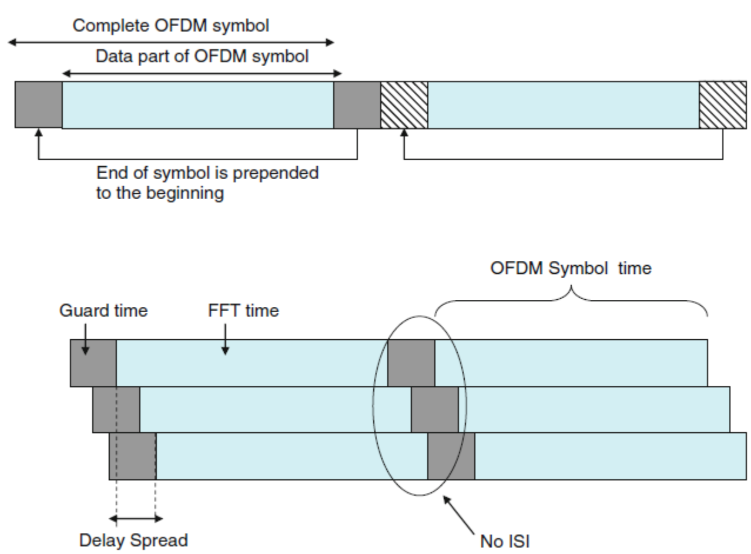


Figure 2.3: illustrations how CP prevents ISI occurrence [4].

2.1.2 OFDM Signal Characteristics

An OFDM signal contains of N orthogonal subcarriers modulated by N parallel data streams, Figure 2.4. After that, a group of block size N, which collected from data symbols ($d_{n,K}$), modulated with complex exponential waveform $\{\emptyset_K(t)\}$. After modulation, the data symbols are transmitted simultaneously as transmitter data stream. The total continuous-time signal consisting of OFDM block is given by

$$x(t) = \sum_{n=-\infty}^{\infty}\left[\sum_{k=0}^{N-1} d_{n,k}\varphi_k(t - nT_d)\right], \tag{2.1}$$

where, $\emptyset_k(t)$ represents each baseband subcarrier and is given by

$$\varphi_K(t) = \begin{cases} e^{j2\pi f_k t} & t\epsilon[0, T_d] \\ 0 & otherwise \end{cases} \tag{2.2}$$

where $d_{n,k}$ is the symbol transmitted during n^{th} timing interval using K^{th} subcarrier, T_d is the symbol duration, N is the number of OFDM subcarriers and f_k is k^{th} subcarrier frequency, which is calculated as $f_k = f_0 + \frac{k}{T}$, $k = 0 \ldots N-1$. Note that f_0 is the lowest used frequency [4].

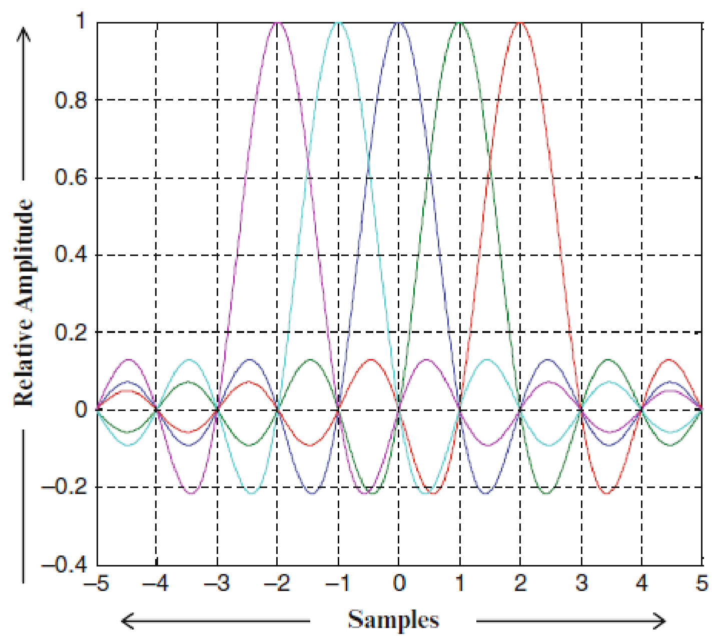


Figure 2.4-a: Spectra for OFDM subcarriers [13].

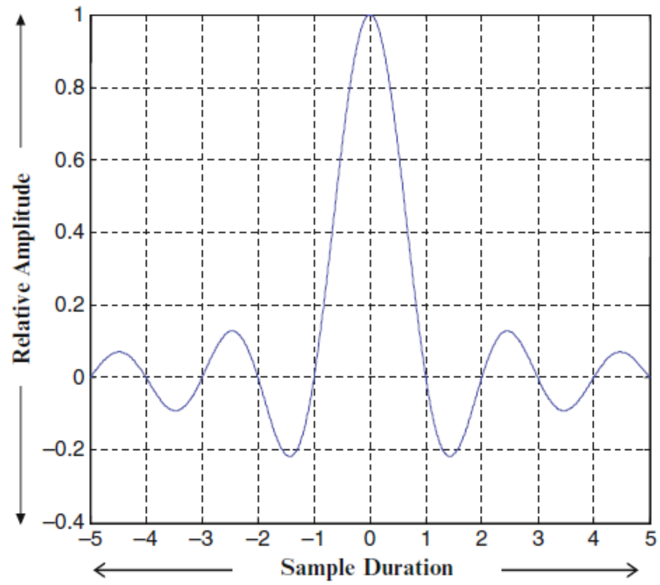


Figure 2.4-b: Spectra for one sample duration [13].

2.1.3 OFDM Transmission

The OFDM integrated to transmit multiple data symbols simultaneously using orthogonal subcarriers which they are modulated using some modulation scheme (QAM, PSK, etc). Thus, the OFDM system is to decompose the high data stream of bandwidth W into N lower rate data streams and then to transmit them simultaneously over a large number of subcarriers that done digitally using the inverse discrete Fourier transform (IDFT). Value of N is kept sufficiently high to make the individual bandwidth (W/N) of subcarriers narrower than the coherence bandwidth (B_c) of the channel. So, the effect of the channel on each subcarrier appears as flat fading. These subcarriers are orthogonal to each other which allows for the overlapping of the subcarriers with the orthogonality ensures the separation of subcarriers at the receiver side. As compared to FDMA systems, which do not allow spectral overlapping of carriers, OFDM systems are more spectrally efficient [4]. As shown in in figure 2.5 [5], the data constellation modulated and demodulated in The IDFT and the DFT, on the orthogonal Subcarriers (SCs). These signal processing algorithms instead of I/Q-modulators and demodulators where the modulation scheme is a mapping of data words to a real (In phase) and imaginary (Quadrature) constellation, also known as an IQ constellation. For example, 64-QAM that would otherwise be required. After that, at the input of the IDFT, N data constellation points $\{x_{i,k}\}$ are present, where N is the number of DFT points. (i is an index on the SC; k is an index on the OFDM symbol). These constellations can be taken according to I/Q modulation ((PSK) or QAM) signaling set (symbol mapping). The *N* output samples of the IDFT, being in TD, form the baseband signal carrying the data symbols on a set of *N* orthogonal SCs. In a real system, however, not all of these N possible SCs can be used for data [5]. After the IDFT operation, a cyclic prefix is added to the OFDM symbol prior to digital- to-analogue converter (DAC). The DAC output is a baseband analog signal which is then up-converted in frequency and transmitted. At the receiver, the received signal is down-converted to baseband. Then the signal is converted from analog to digital using an analogue-to digital converter (ADC). After removing guard interval, the samples are fed into the discrete Fourier transform (DFT) to be converted to frequency domain. Finally, the data is detected.

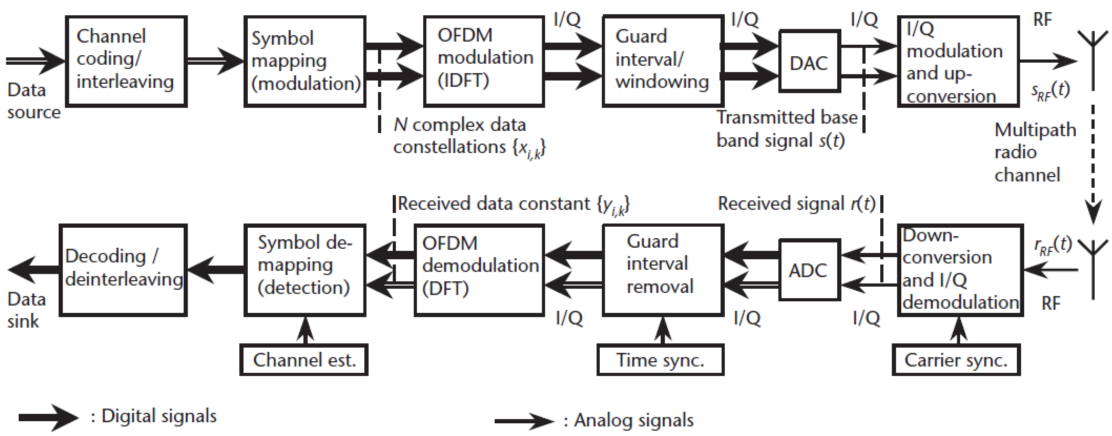


Figure 2.5: OFDM transceiver system [5].

2.1.4 Serial to Parallel Conversion

The data allocated to each symbol depends on the modulation scheme used and the number of subcarriers so the data need to be converted from serial bit stream to the data to be transmitted in each OFDM symbol. For this purpose, serial to parallel conversion block is needed. For example, in case of a subcarrier modulation of 16-QAM, each subcarrier carries 4 bits of data, and so for a transmission using 100 subcarriers the number of bits per symbol would be 400 [4] while the number of parallel symbols entering the IFFT block is 100. The factor for deciding the type of constellation to be used is characteristics of the channel. Thus, in a channel with high interference, a small constellation like BPSK is favorable as the required signal-to-noise-ratio (SNR) in the receiver is low. For interference free channel a larger constellation is more beneficial due to the higher bit rate. Known pilot symbols mapped with known mapping schemes can be inserted at this moment.

2.1.5 Discrete Fourier Transform Implementation

One of the characteristics of OFDM is the orthogonality of subcarriers, SCs, so the signal is transmitted in the time domain. In order to do that, IFFT is used to modulate the data into the time domain and the FFT is used to recover the original data in the receiver. The IFFT allows an efficient implementation of modulation of data onto multiple carriers [4]. Due to the similarity between the forward and inverse transform, the same circuitry, with trivial modifications, can be used for both modulation and demodulation in a transceiver. After that, the flat channel effect on each SC, this leads to an extremely complex architecture involving many oscillators and filters at both transmit and receive ends. Weinsten and Ebert first revealed that OFDM modulation/demodulation can be implemented by using inverse discrete Fourier transform (IDFT)/discrete Fourier transform (DFT) [6].

Let $\{X_n\}^{N-1}{}_{n=0}$ be the complex symbols to be transmitted using OFDM. The modulated OFDM signal can be expressed in the time domain as

$$x(t) = \sum_{n=0}^{N-1} X_n e^{-j2\pi f_n t} \qquad 2.3$$

where $f_n = f_0 + n\Delta f, n = 1, 2 \dots N-1$ t_s *and* Δf are called the symbol duration and the subchannel spacing, respectively, N is the number of subcarriers and k is the period of IFFT. Δf is usually chosen to make the subcarriers orthogonal.

If $x(t)$ is sampled at an interval of $T_{sa} = \frac{T_s}{N}$, then:

$$x_k = x(KT_{sa}) = \sum_{n=0}^{N=1} X_n \, e^{j2\pi f_n \frac{kT_s}{N}} \qquad 2.4$$

Without loss of generality, setting f_o=0 , then $T_s.f_n$=n, then

$$x_k = \sum_{n=0}^{N-1} X_n e^{j2\pi \frac{kn}{N}} = IDFT\{X_n\} \qquad 2.5$$

where, IDFT denotes the inverse Discrete Fourier Transform. Therefore, the OFDM transmitter can be implemented using the IDFT. For the same reason, the receiver can be also implemented using DFT.

The major advantages of DFT/IDFT implementation of OFDM. First, because of the existence of an efficient IFFT/FFT algorithm, the number of complex multiplications for IFFT and FFT is reduced from N^2 to $(N/2)\log2(N)$ [7]. This means that the complexity is linear with the number of subcarriers, N. The other advantage is the simplification of architecture for OFDM implementation when large numbers of subcarriers are required because no need to complex RF oscillators and filters.

2.1.6 Advantages of OFDM

OFDM has more than a few advantages such as high data rate in mobile wireless channel. OFDM also has good avoidance to inter symbol interference and inter carrier interference. The advantages of OFDM are shown below [4]:

- Equalization is very simple compared to Single-Carrier systems.
- OFDM is more robust to frequency selective fading than single carrier system
- OFDM can also support dynamic packet access.
- Smart antennas can be integrated with OFDM.
- Robust against intersymbol interference (ISI) and fading caused by multipath propagation.
- High spectral efficiency compared to non orthogonal multi-carrier like FDMA.
- Efficient implementation using FFT.

2.1.7 Disadvantages of OFDM Systems

The following disadvantages of OFDM may be identified:

- Strict Synchronization Requirement because OFDM is highly sensitive to time and frequency synchronization errors. Thus, any synchronization error can lead to a high bit error rate. There are two sources of synchronization errors: first one is caused by the difference between local oscillator frequencies in transmitter and receiver. While the other is due to the relative motion between the transmitter and receiver that gives Doppler spread.
- Peak to Average Power Ratio (PAPR) is proportional to the number of subcarriers used for OFDM systems where 10 log (N) if N is the number of subcarriers. And so, OFDM system with large number of sub-carriers will thus have a very large that makes the implementation of Digital-to-Analog Converter (DAC) and Analog-to-Digital Converter (ADC) to be extremely difficult. Therefore, there are 3 techniques for reducing PAPR, they are Signal Distortion Techniques, Coding Techniques and finally the Scrambling Technique.
- Using OFDM in cellular systems will give a rise to Co-Channel Interference (CCI) that because the adjacent subcarriers will fall within the coherence bandwidth and will thereby experience flat fading. So, CCI is combated by combining adaptive antenna techniques, such as sectorization, directive antenna, antenna arrays, etc.

2.2 Long-Term Evolution LTE

2.2.1 LTE Overview

Long Term Evolution or LTE, which was introduced in 3GPP Release 8, is the next major step in mobile radio communications. It will give an experience for users to interact with applications, such as HD TV broadcasting, user-generated videos, advanced games, and professional services. LTE uses OFDM radio access technology together with advanced antenna technologies. In addition, LTE is a versatile technology that fulfills or exceeds 3GPP requirements [8].

LTE is developed for a number of frequency bands – E-UTRA operating bands- currently ranging from 700 MHz up to 2.7 GHz, downlink peak rates of more than 100 Mbps and roundtrip time in the Radio Access Network (RAN) of less than 10ms. LTE supports flexible carrier bandwidths from less than 1.2 MHz up to 20 MHz in many new and existing spectrum bands, and adopts time division duplex (TDD) or frequency division duplex (FDD) deployments. Additionally, LTE supports the handover and roaming to existing mobile networks, thereby providing ubiquitous coverage to all mobile subscribers from the very outset [8].

2.2.2 Overall LTE Architectural Overview

The LTE is a growth of the radio and the non-radio access [9]. The radio access basically is the evolution of the LTE Physical Layer, while the non-radio access grouped under the System Architecture Evolution (SAE). The major components of the LTE System Architecture are:

1. User Equipment (UE)
2. Radio Access Network (RAN)
3. Evolved Packet Core (EPC)

Figure 2.6 describes the architecture and network elements in the architecture configuration of LTE system. The logical nodes and connections shown in this figure represent the basic system architecture configuration. These elements and functions are needed in all cases when E-UTRAN is involved. This figure also shows the division of the architecture into four main high level domains: User Equipment (UE), Evolved UTRAN (E-UTRAN), Evolved Packet Core Network (EPC), and the Services domain.

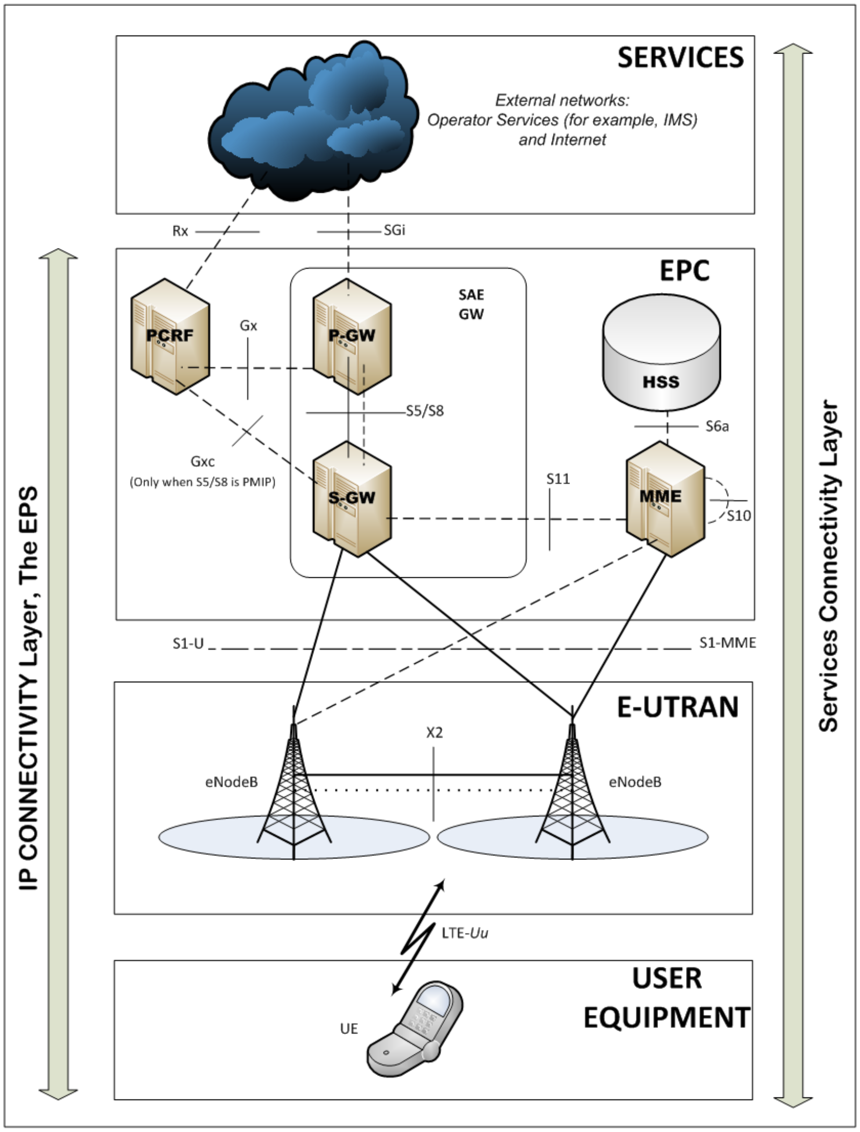


Figure 2.6: Architecture for LTE access networks [17].

LTE Radio Access Network and Evolved Packet Core (EPC) are comprised to The Evolved Packet System (EPS) (EPS >> RAN + EPC). Furthermore, The UE, EPC and the E-UTRAN are the integral parts that form the Internet Protocol Connectivity Layer, which is also referred to the EPS. The EPS provides the IP based connectivity services, with all services offered at the top of the IP layer [10].The mean part we focused on is the physical layer, so the radio transmissions in LTE are based on the Orthogonal Frequency Division Multiplexing (OFDM) modulation scheme. The OFDM is used in downlink transmissions and the other scheme for uplink transmissions is the Single Carrier Frequency Division Multiple Access (SC-FDMA). SC-FDMA allow multiple access by assigning sets of sub-carriers to each individual users. Moreover, OFDMA can exploit subsets of sub-carriers distributed inside the entire spectrum whereas SC-FDMA can use only adjacent subcarriers. OFDMA is able to provide high scalability, simple equalization, and high robustness against the time-frequency selective fading of the radio channel. On the other hand, SC-FDMA is used in the LTE uplink to increase the power efficiency of user equipment (UEs) which are battery supplied [11].

2.2.3 Core Network

The Core Network (CN), also known as EPC, responsible of overall control of the UE and establishes the bearers which the bearers refer to the IP packet flow that defines the Quality of Service (QoS) between the User Equipment (UE) and the Gateway (GW). The core network has a number of different logic nodes, some of which are [9]:

1. Mobility Management Entity (MME).
2. Serving Gateway (S-GW).
3. Packet Data Network (PDN) Gateway (P-GW).
4. Policy and Charging Rules Function (PCRF).
5. Home Subscriber Service (HSS).

1. Mobility Management Entity, MME:

   The MME is the main control node in the EPC. The control plane information coming from the eNodeB is mainly routed to the MME. Thereby the MME also has a logically direct CP connection to the UE and one of the most essential functions of the MME is that it handles the signaling between the UE and the CN and this connection is used as the primary control channel between the UE and the network. Also, it handles the issue of security and authentication for keys offering; in addition to mobility management - where the MME does management functions by making request setup and release of appropriate resources in eNodeB and the S-GW; the MME also manages the subscription protocol and service connectivity. The responsible protocols between the UE and the CN are the Non-Access Stratum (NAS) protocols. Also, MME has mobility management function where it keeps track of the location of all UEs in its service area. When a UE makes its first registration to the network, the MME will create an entry for the UE, and signal the location to the home subscriber server (HSS) in the UE's home network, figure 2.7 shows mean functions of MME [9].

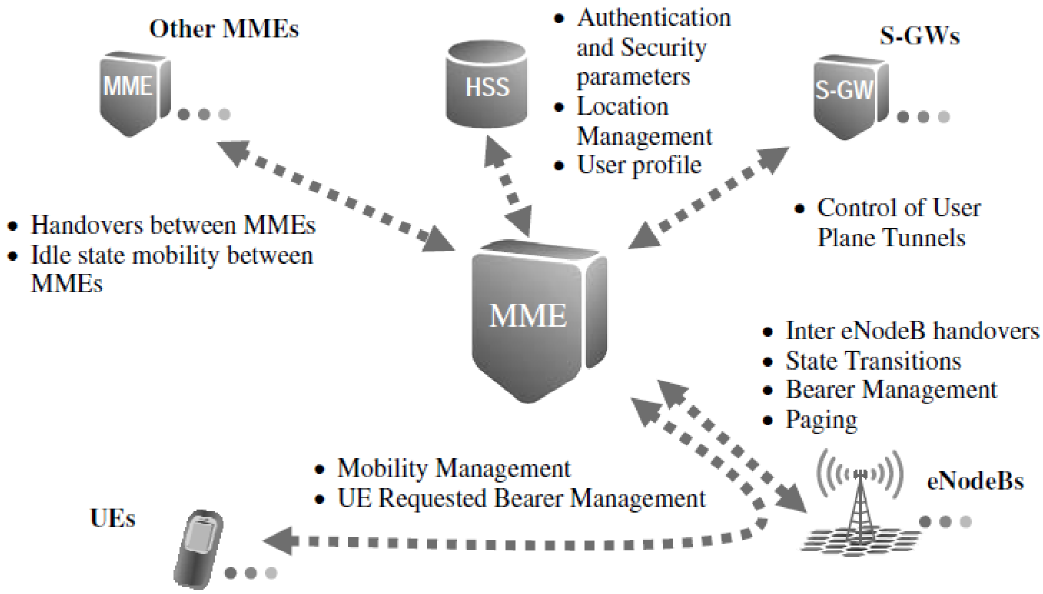


Figure 2.7: MME connections to other logical nodes and main functions [17].

2. Serving Gateway (S-GW)

S-GW is responsible for the Up-plane tunnel management and switching; it acts as the mobile anchor between EPC and the LTE RAN. All the user's packets are routed through the S-GW. Although, the S-GW has a role in control functions, it is very important in terms of inter-connectivity to other 3GPP technologies like GPRS/GSM and UMTS. Also, when the UEs' bearers are setup, cleared or undergo modification, the S-GW make resource allocation depending on the various requests from the MME, P-GW and/or PCRF. During mobility between eNodeBs, the S-GW acts as the local mobility anchor. The MME response of handover of UE data from source eNodeB to target eNodeB. The mobility scenarios also include changing from one S-GW to another, and the MME controls this change accordingly, by removing tunnels in the old S-GW and setting them up in a new S-GW [9].

3. Packet Data Network (PDN) Gateway (P-GW)

The P-GW serves as the end point intermediary router between the EPS and external networks. It mainly provides IP connection at its active point; and is refer to as the highest level mobility or final anchor in the system. Also, it does IP addressing to UEs, performs traffic gating and filtering duties when needed [9].

4. Policy and Charging Rules Function (PCRF)

   The PCRF is responsible for the QoS as well as the policy control decision making. Also, it controls the flow-based charging for functions within the Policy Control Enforcement Function (PCEF) which is part of the P-GW. In other words, it does the Policy and Charging Control (PCC) functions. Therefore, PCRF is a server usually located with other CN elements in operator switching centers [9].

5. Home Subscription Server (HSS)

   The HSS is a database that contains all user's subscription details. It contains the information about the PDN every user is connected to or can connect to. Essentially, it holds all permanent subscribers' data. It also records the location of the user in the level of visited network control node, such as MME. Furthermore, the HSS also stores the Identities of those P-GWs that are in use. As part of its side functions, the HSS can also integrate the Authentication Centre (AuC) [9].

2.2.4 LTE Physical Layer

OFDMA at the physical layer, in combination with a Medium Access Control (MAC) layer, provides an optimized resource allocation and Quality of Service (QoS) support for different types of services. The data transported to the higher layers through physical layer which services transport via the MAC sub-layer. The physical layer is defined in a bandwidth adaptation way. The main functions of physical layer include the following [12]

- Transport channel error detection and report to the higher layers
- FEC encoding and decoding
- Transport channel rate adaption to the physical channel
- Transport channel mapping onto the physical channel
- Physical channel modulation/demodulation
- Synchronization of time and frequency
- Reporting radio channel measurements to higher layers
- MIMO antenna signals processing, transmit diversity, and beam forming

Physical layer in LTE is supporting both FDD (Frequency Division Duplex) and TDD (Time Division Duplex) frame structure where FDD separated downlink and uplink in frequency domain and which TDD separated downlink and uplink in time domain. Both FDD and TDD share the same framing structure. This frame has duration of 10 ms and consists of 20 time slots. A sub-frame is formed by two adjacent time slots and it spans to 1ms (i.e. 0.5 ms x 2) [12], where there are additional framing for TDD. The following points show Downlink physical channels where they control the transmission of user data and control information from the eNodeB towards UE.

- PDSCH (Physical Downlink Shared Channel)
- PDCCH (Physical Downlink Control Channel)
- CCPCH (Common Control Physical Channel)

Additionally, the Uplink physical channels include the following channels

- PUSCH (Physical Uplink Shared Channel)
- PUCCH (Physical Uplink Control Channel)

LTE downlink/uplink channels used QPSK, 16QAM, or 64QAM and the broadcast channel have use only QPSK modulation.

Multiple Input Multiple Output (MIMO) transmission technique has been specified for LTE downlink deployment in which antenna configuration is represented by N x M where N (1-4) is the number of transmit antennas and M (1-2) is the number of receive antennas. MIMO has been specified to exploit the multipath fading so that spatially multiplexed independent data streams can be transmitted. MIMO implementation will get impaired in a low multipath distortion environment [12].

As we mention two radio frame structures are supported. Type 1, applicable to FDD and type 2, applicable to TDD.

1. Frame structure type 1

   Frame structure type 1 is appropriate to both full duplex and half duplex FDD. The radio frame lasting, Tf, 10ms long and consists of 20 slots of length T_{slot} = 0.5ms which numbered from 0 to 19. A subframe is defined as two consecutive slots where subframe, i consists of slots 2i and 2i +1. For FDD, 10 subframes are available for downlink transmission and 10 subframes are available for uplink transmissions in each 10 ms interval. Uplink and downlink transmissions are separated in the frequency domain. In half-duplex FDD operation, the UE cannot transmit and receive at the same time while there are no such restrictions in full-duplex FDD, that shows in figure 2.8 [13].

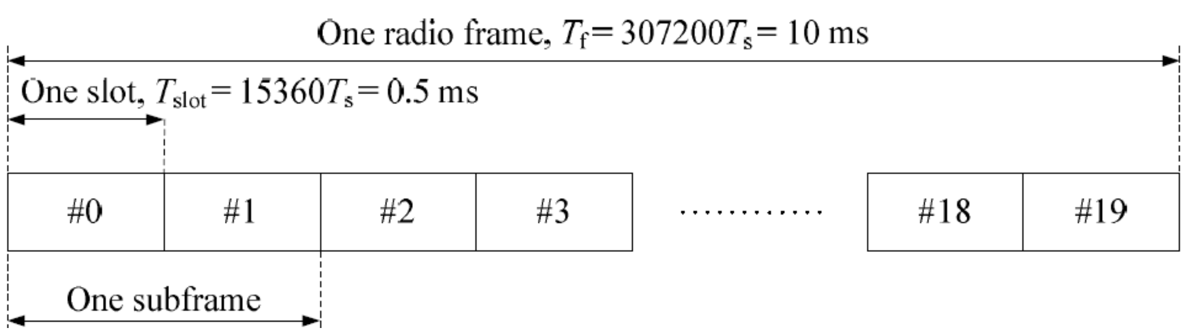


Figure 2.8: FDD frame structure [13].

2. Frame structure type 2

   Frame structure type 2 is applicable to TDD. Frame length T_f lasting10ms consists of two half frames of length 5ms. Each half-frame consists of five subframes of length 1ms .Thus, subframe in a radio frame, '*D*' denotes the subframe is reserved for downlink transmissions, 'U' denotes the subframe is reserved for uplink transmissions and 'S' denotes a special. Each subframe, i, is defined as two slots, 2i and 2i +1 of length T_{slot} 0.5

ms in each subframe. Uplink-downlink configurations with two time frames (5 ms and 10 ms). Which, In case of 5 ms downlink-to-uplink switch-point periodicity, the special subframe exists in both half-frames. In case of 10 ms downlink-to-uplink switch-point periodicity, the special subframe exists in the first half-frame only. Subframes 0 and 5 and DwPTS are always reserved for downlink transmission. UpPTS and the subframe immediately following the special subframe are always reserved for uplink transmission see figure 2.9 [13].

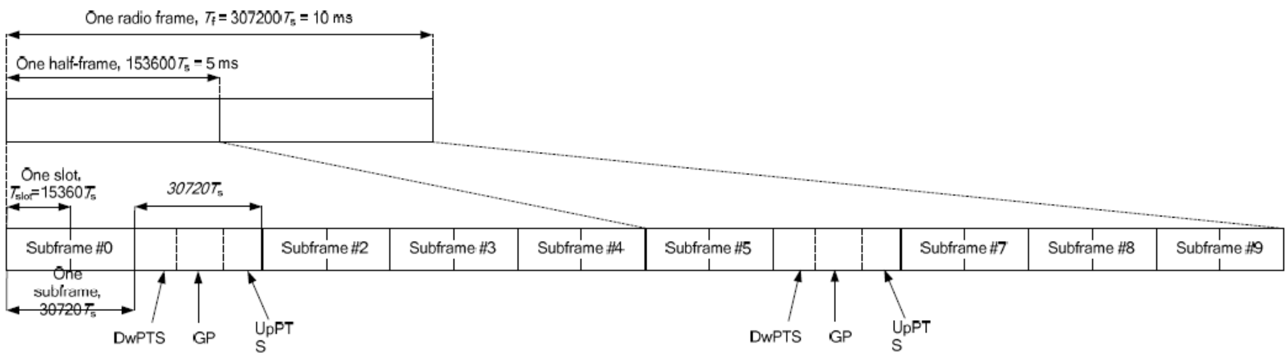


Figure 2.9: TDD frame structure [13].

2.2.5 Physical Resource and Slot structure

Recourse gird or element is smallest unit that is corresponding to one OFDM subcarrier during OFDM symbol interval. The number of sub-carriers is being determined by the transmission bandwidth. For normal cyclic prefix (CP) each slot contains seven OFDM symbols and in case of extended cyclic prefix, 6 OFDM symbols are slotted-in in each time slot this mentioned in section 2.2.1. Thus, the mapping of physical channel of resource element is called resource block [13].

As we mention, in LTE downlink a constant sub-carriers spacing of 15 kHz is utilized. In frequency domain, 12 sub-carriers are grouped together to form a Resource Block (RB) occupying total 180 kHz in one slot duration as illustrated in figure 11 wherever, In case of short CP, length a resource block contains 84 resource elements (RE) and for long CP the number of RE is 74. Figure 1.10 shows the LTE RB grid approach illustration with both time domain and frequency domain diagrams as indicated and table 1 illustrates the number of resource blocks, RB, in each bandwidth and cyclic prefix [12].

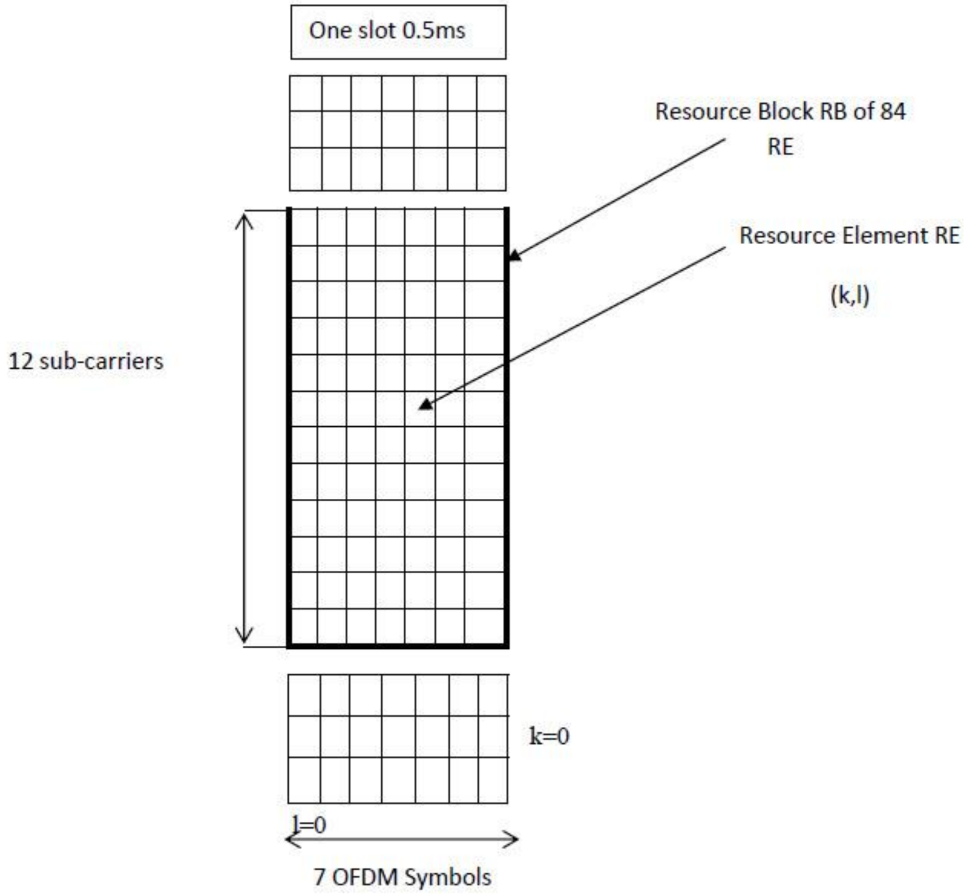


Figure 2.10: Downlink Resource Block [18].

Bandwidth (MHz)	1.25	3	5	10	15	20
Subcarrier bandwidth (KHz)	15					
Physical resource block (PRB) bandwidth (KHz)	180					
Number of available PRBs	6	15	25	50	75	100

Table 1: relation of resource block to bandwidth [12].

2.2.6 Sub-Carrier Allocation and User Scheduling

The OFDMA characterizes every user with a group of sub carriers with specific time slot in other word resource block. In a single channel spectrum several sub carriers are allocated to every user based on his requirement. This scheme is particularly useful in downlink when the number of users are high. If the data rate required is low then scheme is adapted as it consumes less resource and the delay is reduced effectively. The mobile users can be synchronized in time domain and frequency domain. This makes the uplink to be orthogonal and in sync [13].

2.3 Femtocell

2.3.1 Femtocell Overview

Femtocells, also known as ‘home base station’, are cellular network access points that connect to user equipment UE to a mobile operator’s network via residential DSL, optical fibres or wireless last-mile technologies. Thus, it installed by end user and it called home-eNB (HeNB) for LTE network system. Moreover, it is a part of self organizing network (SON) with zero touch installation. It supports limited number of active connections 3 or 4 but can be extended from 8 to 32. Also, it is used to improve indoor signal strength as it avoids walls penetration loss. The main driver for user is the improved coverage and capacity thus offers better quality of service, not only to indoor users but also to outdoor users by offloading. Femtocell uses the licensed spectrum owned by a mobile operator. Smaller cells are typically used in homes and there is an option for enterprise femtocells. Figure 2.11 shows a simple femtocell implementation [14].

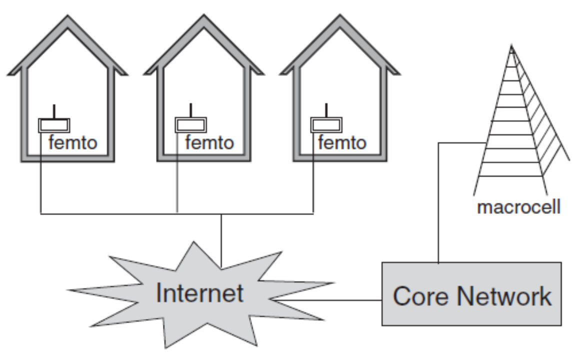


Figure 2.11: simple femtocell implementation scenario [14].

2.3.2 Femtocells features and attributes

The following points characterize femtocell technology, among these features [15]

1. Usage of mobile technology: the target of femtocell technology complementing the current cellular systems. So, it integrated with the readily available cellular components including mobile phones and cellular protocols and interfaces like as GSM, WCDMA, LTE, Mobile WiMAX and CDMA. In order that, it standardized by 3GPP, 3GPP2 and the IEEE/WiMAX forum.
2. Operation in licensed spectrum: This allows supporting regular mobile device without the need of dual-mode devices to be used in a femtocell.
3. Coverage and capacity enhancement: FBSs supports region which is not covered by Macro Base Station MBS (M-eNB in LTE system) so femtocell enhanced this problem. Moreover, it transmits in relatively very low power targeting a very small indoor area compared to Macrocell Base Stations (MBSs). The very short distance between a transmitter and a receiver promotes the use of higher order modulation schemes and hence promotes higher capacity.
4. Backhauled to the cellular network: Data sent from a FBS is backhauled to the cellular network through DSL or cable, using standard Internet protocols.
5. Simplification of installation: FBSs will support plug and play this features don’t need from end users to have any complex information about configuration. Thus, femtocell initiates configuration and can re-configure and heal itself without any touch from external user. Although the FBSs will receive their operation parameters via the operator network, this operation will be done automatically under the hood, without any user intervention.

2.3.3 Advantages of femtocells

Femtocells can bring a lot of advantages for both operators and subscribers [14] [15].

- Operator’s Perspective

The benefits of employing femtocells to the operator as follow

- Increase system coverage and capacity: By offloading indoor connections to a femtocell instead of a macrocell. This frees up resources on the macrocell to service more outdoor users. Also, this increases the capacity of indoor users by connecting them to a near indoor FBSs.
- Filling coverage holes: Femtocells will cover a large portion of the gaps. This gaps because the MBSs adjust its transmission power to mitigate the interference in the overlapped area between two MBSs, which results in the creation of pockets' or `holes' between macrocells. Unfortunate users who reside within these holes experience very low signal level, lower than what may be required to make a call. So, femtocell perfectly in situations like this. Minimal interference from macrocells due to their poor signals will hardly affect the femtocell performance.
- Reduce churn: Poor indoor coverage can cause churn, where Churn represents a considerable loss to operators especially in saturated markets. However, with the introduction of femtocells, value-added family packages can be delivered to the customer, and hence, reducing churn to a great extent.
- Reduction costs: By increasing the system capacity without the need for new cell sites for macrocells, the operator will extremely savings CAPEX and OPEX.

- Subscriber’s Perspective

This section presents some of attractions that femtocells can offer to subscribers.

- Superior service: Moreover to the advantages stated before, femtocells will also support the operator provider better-off services like femto zone-based services, bundled services and other bandwidth consuming application, which will also encourage the customer use them indoors.
- Saving power: Because of the short distance between a FBS and a UE; battery operated devices will communicate using lower power levels than those required to communicate to a macrocell, resulting in longer battery life. This will also reduce health concerns on using mobile devices.

- Centralized management: Femtocells can offer a single address book and one billing account for both land line phone, broadband and mobile phone.
- Cheaper prices: By installing new base stations inside homes, that will reduce the cost in provider that leads to decrease of the service cost and offer value-added services like free or cheap calls from home.

2.3.4 Femtocell applications

The followings are some of the applications of femtocell [14] [15]

- Femtozone services: In addition of the services based on the presence of mobile devices, like calls and sending SMS, femtocell can notification the user when he enters the femtozone, or synchronize pictures and videos from a trip. Another hopeful application can be a `virtual number' to reach all people currently in the house, or for holding conference calls with family members. Enterprise VoIP and file transfer can also be categorized as femtozone services.
- Connected home services: Femtocell affective tool will service the smart device in smart home that is able to control home equipments via a mobile.

2.3.5 The Access Method

Femtocells can be configured in three ways:

• Open access: all users are allowed to connect.

• Closed access: the femtocell allows only specific subscribed users to establish connections.

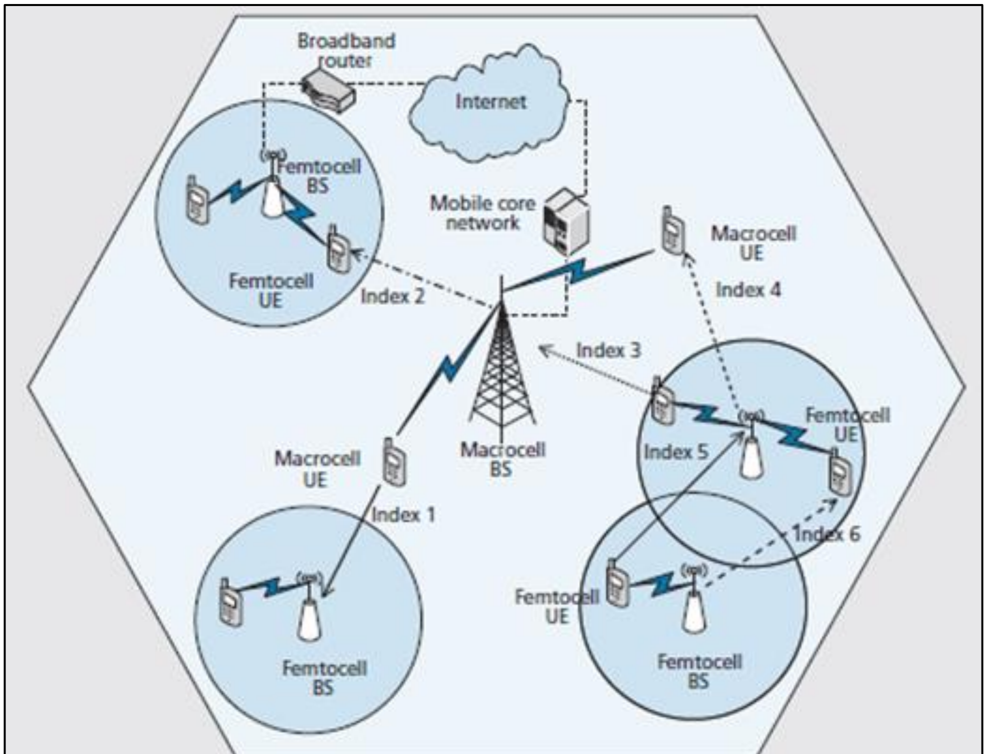
• Hybrid access: nonsubscribers use only a limited amount of the femtocell resources [16].

2.3.6 Technical Challenges in Femtocell Deployment

The mass deployment of femtocells gives rise to several technical challenges as follow [16].

- Interference management: One of the major challenges is interference management between neighboring femtocells and between femtocell and macrocell. In general, the first one is Co-tier interference, this type of interference occurs among network elements that belong to the same tier in the network, in this co-tier interference occurs between neighboring femtocells. However, in OFDMA systems, the co-tier uplink or downlink interference occurs only when the attacker and the victim use the same sub-channels. Therefore, efficient allocation of sub-channels is required in OFDMA-based femtocell networks to mitigate co-tier interference. The other interference is Cross-tier interference where this type occurs among network elements that belong to the different tiers of the network which femtocell UEs and macrocell UEs act in same sub carrier. Again, in OFDMA-based femtocell networks, cross-tier uplink or downlink interference occurs only when the same sub-channels are used by the aggressor and the victim. Due to spectrum insufficiency, the femtocells and macrocells have to reuse and/or share the total allocated frequency band partially or totally which leads to cross-tier or co-channel interference, two types shown in figure 2.12.
- Handoff and mobility management: An effective and efficient mobility management and handover scheme (macrocell-to femtocell, femtocell-to-macrocell and femtocell-to-femtocell) is necessary for mass deployment of femtocells in UMTS and LTE networks. The scheme should have low complexity and signaling cost, deal with different access modes and perform proper resource management beforehand for efficient handover.

Timing and synchronization is one of the major challenges for femtocells since synchronization over IP backhaul is difficult, and inconsistent delays may occur due to varying traffic congestion. Since the femtocells are required to operate on a “plug-and-play” basis, it is important that femtocells can organize and configure autonomously and access the radio network intelligently so that they only cause minimal impact on the existing macrocell network.



Index	Aggressor	Victim	Interference type	Transmission mode	Symbol
1	Macrocell UE	Femtocell BS	Cross-tier	Uplink	——————>
2	Macrocell BS	Femtocell UE	Cross-tier	Downlink	—·—·—·—>
3	Femtocell UE	Macrocell BS	Cross-tier	Uplink	———————>
4	Femtocell BS	Macrocell UE	Cross-tier	Downlink	--------->
5	Femtocell UE	Femtocell BS	Co-tier	Uplink	——————>
6	Femtocell BS	Femtocell UE	Co-tier	Downlink	------>

Figure 2.12 Interference scenarios in femtocell networks [16].

2.4 Channel Modelling

2.4.1 Radio Propagation Models

The demand for realistic mobile fading channels is shown in the beginning of wireless communications. The reason for this importance is that efficient channel models are essential for the analysis, design, and deployment of communication system for reliable transfer of information between two nodes. Accurate channel models are also significant for testing, parameter optimization and performance evolution of communication systems. So, the performance and complexity of signal processing algorithms, transceiver designs and smart antennas etc. The difficulties in modeling a wireless channel are due to complex propagation processes. A transmitted signal arrives at the receiver through different propagation mechanisms shown in figure 2.13. The propagation mechanisms involve the following basic mechanisms:

- Free space or line of sight propagation.
- Specular reflection due to interaction of electromagnetic waves with plane and smooth surfaces which have large dimensions as compared to the wavelength of interacting electromagnetic waves.
- Diffraction caused by bending of electromagnetic waves around corners of buildings.
- Diffusion or scattering due to contacts with objects having irregular surfaces or shapes with sizes of the order of wavelength.
- Transmission through objects which cause partial absorption of energy.

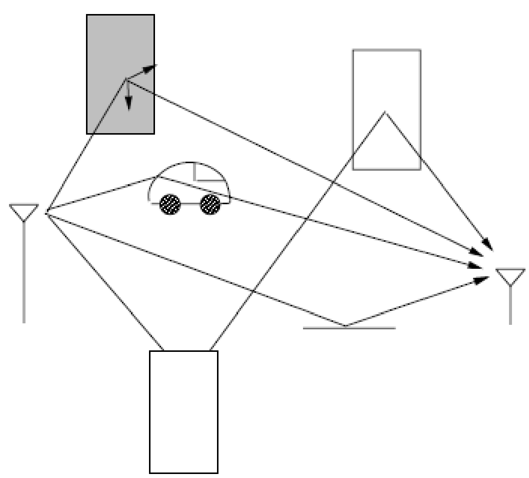


Figure 2.13: Signal propagation through different paths [17].

To predict the performance of narrowband receivers, classical channel models which provide information about signal power level distributions and Doppler shifts of the received signals, may be sufficient. The advanced technologies, (e.g., UMTS and LTE) build on the typical understanding of Doppler spread and fading; also incorporate new concepts such as time delay spread, direction of departures (DOD), direction of arrivals (DOA) and adaptive array antenna geometry. The presence of multipath (multiple scattered paths) with different delays, attenuations, DOD and DOA gives rise to highly complex multipath propagation channel [17].

2.4.2 Multipath Propagation Channel

The different propagation path lengths cause different propagation time delays that received in UE. The power distribution of channel taps is described by a distribution function depending on the propagation environment. Thus, power delay profile (PDP) of a multipath channel depending on the phases the multipath signals interact to environment. The most severe multipath channel is Rayleigh fading channel in which there is no line of sight path and the channel taps are independent. In the case of Rician fading channel the fading dips are low due to the presence of line of sight component in addition to the dispersed paths. A radio channel can be characterized as narrowband or wideband channel where it depending on the channel characteristics and duration of a symbol.

Narrowband system is the system which the time difference between the first and the last pulse is smaller than the duration of a symbol if the impulse response of a multipath channel consists of a series of pulses spread in time or frequency and if the time difference between the first and last pulse is greater than the duration of a symbol, then the system is called wideband system. Wideband characteristics of impulse response are often used to describe the behavior of the multipath channel. The degradation in the received signal level due to multipath effects can be classified into large scale path loss component and small scale fast fading component with Rayleigh or Rician distribution depending on the absence or presence of LOS component between the transmitter and receiver [17]. Thus, a propagation model can be used to describe a wireless cellular environment as follow.

1. Large Scale Propagation Model: Large scale propagation model is used to characterize the received signal strength by averaging the amplitude or power level of the received signal over large transmitter-receiver separation distances in the range of hundreds or thousands of a wavelength. The large scale models are often derived from measured data. However, semi-empirical models are employed in smaller areas to achieve higher accuracy. The theoretical models are used which are then fitted to measured data to obtain desired model for a particular propagation scenario [17].
2. Small Scale Propagation Model: the three most important effects of small-scale fading effects are:

   • Rapid changes in signal strength over a small travel distance or time interval.

   • Random frequency modulation due to varying Doppler shifts on different multipath signals.

   • Time dispersion (echoes) caused by multipath propagation delays.

This model is used to characterize the rapid variations of the received signal strength due to changes in phases when a mobile terminal moves over small distances on the order of a few wavelengths or over short time durations on the order of seconds. Since the mean power remains constant over these small distances, small scale fading can be considered as superimposed on large scale fading for large scale models. The most common description of small scale fading is by means of Rayleigh distribution [17].

The behavior of a multipath channel needs to be characterized in order to model the channel. The concepts of Doppler spread, coherence time, and delay spread and coherence bandwidth are used describe various aspects of the multipath channel. Delay spread and coherence bandwidth are parameters which describe the time dispersive nature of the channel in a local area. However, they do not offer information about the time varying nature of the channel caused by either relative motion between the mobile and base station, or by movement of objects in the channel. Doppler spread and coherence time are parameters which describe the time varying nature of the channel in a small-scale region.

2.4.3 Delay Spread

The time dispersion or multipath delay spread related to a small scale fading of the channel needs to be calculated in a convenient way. One simple measure of delay spread is the overall extent of path delays called the excess delay spread. This is not convenient way because different channels with the same excess delays can exhibit different power profiles which have more or less impact on the performance of the system under consideration. A more efficient method to determine channel delay spread is the root mean square (rms) delay spread (τ_{rms}) which is a statistical measure and gives the spread of delayed components about the mean value of the channel power delay profile. Mathematically, rms delay spread can be described as second central moment of the channel power delay profile which is written as follows [18]:

$$\tau_{rms} = \sqrt{\frac{\sum_{n=0}^{N-1} p_n\,(\tau_n - \tau_m)^2}{\sum_{n=0}^{N-1} p_n}} \qquad 2.6$$

$$\tau_m = \frac{\sum_{n=0}^{N-1} P_n \tau_n}{\sum_{n=0}^{N-1} P_n} \qquad 2.7$$

where τ_m is the mean excess delay.

2.4.4 Doppler Spread and Coherence Time

The Doppler spread and coherence time are parameters which describe the time varying nature of the channel in a small-scale region. Doppler spread, B_D is a measure of the spectral broadening caused by the time rate of change of the mobile radio channel and is defined as the range of frequencies over which the received Doppler spectrum is essentially non-zero. When a pure sinusoidal tone of frequency f_c is transmitted, the received signal spectrum, called the Doppler spectrum, will have components in the range $f_c - f_d$ to $f_c + f_d$, where f_d is the Doppler shift. The amount of spectral broadening depends on f_d which is a function of the relative velocity of the mobile, and the angle, θ between the direction of motion of the mobile and direction of arrival of the scattered waves. If the baseband signal bandwidth is much greater than B_D, the effects of Doppler spread are negligible at the receiver. This is a slow fading channel.

Coherence time T_0 is the time domain dual of Doppler spread and is used to characterize the time varying nature of the frequency depressiveness of the channel in the time domain. The Doppler spread and coherence time are inversely proportional to one another [18]. That is, $T_c=1/f_m$.
In other words, coherence time is the time duration over which two received signals have a strong potential for amplitude correlation. If the reciprocal bandwidth of the baseband signal is greater than the coherence time of the channel, then the channel will change during the transmission of the baseband message, thus causing distortion at the receiver. If the coherence time is defined as the time over which the time correlation function is above 0.5, then the coherence time is approximately
$T_c=9/(16\pi f_m)$, where f_m is the maximum Doppler shift.

2.4.5 Types of Small-Scale Fading

Depending on the relation between the signal parameters such as bandwidth, symbol period, etc, and the channel parameters such as *rms* delay spread and Doppler spread, different transmitted signals will undergo different types of fading.
The time dispersion and frequency dispersion mechanisms in a mobile radio channel lead to four possible distinct effects, which are manifested depending on the nature of the transmitted signal, the channel, and the velocity. While multipath delay spread leads to time dispersion and frequency selective fading, Doppler spread leads to frequency dispersion and time selective fading. The two propagation mechanisms are independent of one another. Figure 2.14-A/B show a tree of the four different types of fading [18].

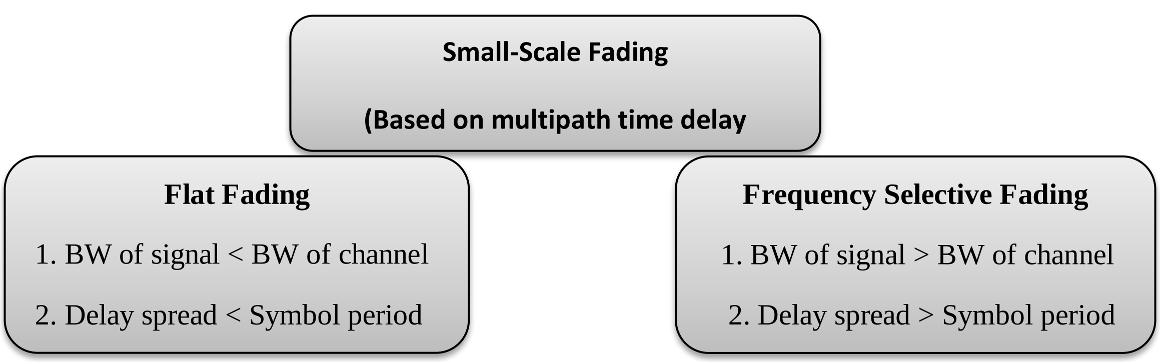


Figure 2.14-A: Types of small-scale fading [18].

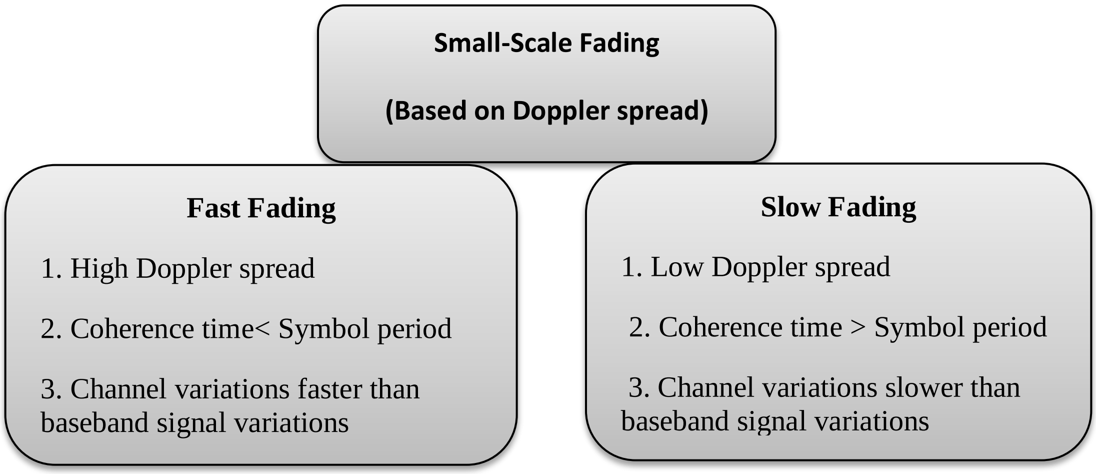


Figure 2.14-A: Types of small-scale fading [18].

Frequency flat fading: A channel is referred to as a frequency flat if the coherence bandwidth $\Delta fc >> B$, where B is the signal bandwidth. All frequency components of the signal will experience the same amount of fading.

Frequency selective fading: A channel is referred to as frequency selective if the coherence bandwidth $\Delta fc \leq B$. In this case different frequency components will undergo different amount of fading. The channel acts as a filter since the channel coherence bandwidth is less than the signal bandwidth; hence frequency selective fading takes place show figure 2.15 [18].

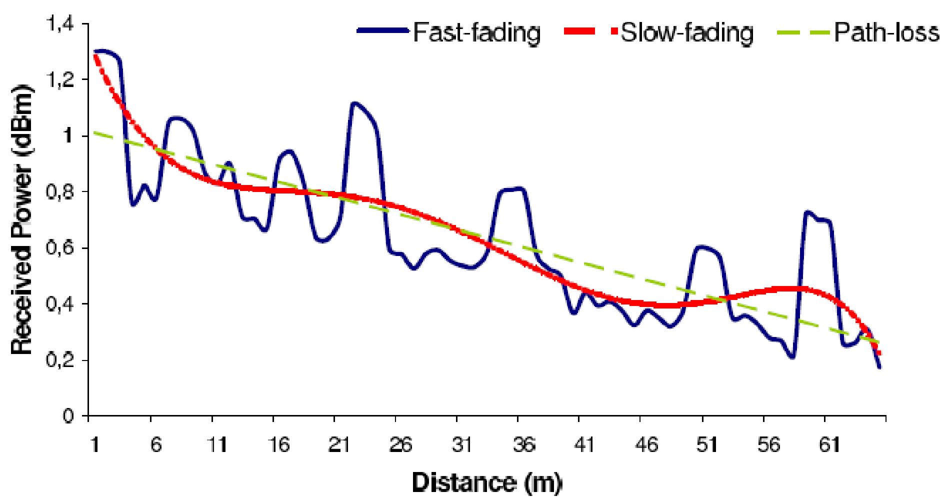


Figure 2.15: A brief description of channel modulation [18].

2.5 Machine Learning

2.5.1 Reinforcement Learning: An Introduction

As we mentioned in previous sections, reinforcement learning is a technique to learn what to do, how to map situations to actions. Thus, reinforcement learning is defined not only by characterizing learning methods, but also by characterizing a learning problem. Clearly, such an agent must be able to sense the state of the environment to some extent and must be able to take actions that affect the state. The agent also must have a goal or goals relating to the state of the environment. The formulation is intended to include just these three aspects, sensation, action, and goal. Reinforcement learning is different from supervised learning which supervised learning is the kind of learning studied in most current research in machine learning if the target, that machine wants to learn, is known like as pattern recognition and artificial neural networks. Moreover, supervised learning is learning from examples provided by a knowledgeable external supervisor. This is an important kind of learning but it is not adequate for learning from interaction with the environment deal with. One of the challenges that arise in reinforcement learning and not in other kinds of learning is the tradeoff between exploration and exploitation [19].

Exploitation is the characteristic to obtain a the pest 'optimum ' action from the experience of agent but that is need to exploration the environment and discover such actions, that has not selected before. The agent has to exploit what it already known in order to obtain reward, but it also has to explore in order to make better action selections in the future. The dilemma is that neither exploration nor exploitation can be pursued exclusively without failing at the task. The agent must try a variety of actions and progressively favor those that appear to be best. On a stochastic task, each action must be tried many times to gain a reliable estimate its expected reward. For now, we simply note that the entire issue of balancing exploration and exploitation does not even arise in supervised learning as it is usually defined.

When reinforcement learning involves planning, it has to address the interplay between planning and real-time action selection, as well as the question of how environmental models are acquired and improved. When reinforcement learning involves supervised learning, it does so for specific reasons that determine which capabilities are critical and which are not. For learning research to make progress, important sub problems have to be isolated and studied, but they should be sub problems that play clear roles in complete, interactive, goal-seeking agents, even if all the details of the complete agent cannot yet be filled in [19].

2.5.2 Elements of Reinforcement Learning

Further than the agent and the environment, the main sub elements of a reinforcement learning (RL) system is a policy, a reward function, a value function, and, optionally, a model of the environment [19].

- A policy defines the learning agent's way of behaving at a given time. In other meaning, policy is a mapping from perceived states of the environment to actions to be taken when in those states. In some cases the policy may be a simple function or lookup table, whereas in others it may involve extensive computation such as a search process.

The policy is the core of a reinforcement learning agent in the sense that it alone is sufficient to determine behavior. In general, policies may be stochastic.

Policies can be classified into two groups.

1) The behavior policy, which determines the conduct of the agent, i.e. the actual action selected by the agent in the current state and.

2) The estimation policy, which determines the policy evaluated, or the action in the next state used for the evaluation of behavior policy. In RL there are two methods to ensure a sufficient exploration, the on-policy and off-policy methods, which basically differ on the form they select the estimation policy.

• On-policy methods: These methods evaluate or improve the policy, symbolizes as π, used to perform the decisions. In other words, the policy followed by the agent to select its behavior in a given state. Thus behavior policy is the same used to select the action.

• Off-policy methods: These methods do distinguish between behavior and estimation policies. Therefore, the policy to generate behavior, π is unrelated to the policy evaluated. The policy evaluated in off-policy methods is the one corresponding to the best action in the next state π^*.

- A reward function defines the objective in a reinforcement learning problem. In other words, it is mapping each perceived state 'or state-action pair' of the environment. A reward indicates the intrinsic desirability of that state. A reinforcement learning agent's goal is to maximize the total reward it receives in the long run. The reward function defines what the good and bad events are for the agent.

In a biological system, it would not be inappropriate to identify rewards with pleasure and pain. They are the immediate and defining features of the problem faced by the agent. As such, the reward function must necessarily be unalterable by the agent. It may, however, serves as a basis for altering the policy. For example, if an action selected by the policy is followed by a low reward, then the policy may be changed to select some other action in that situation in the future. In general, reward functions may be stochastic.

- A value function is the value of a state which the total amount of reward an agent can expect to accumulate over the future. Whereas, the rewards determine the immediate, intrinsic desirability of environmental states, values indicate the long-term desirability of states after taking into account the states. Whereas values correspond to a more refined and farsighted judgment of how pleased or displeased we are that our environment is in a particular state. Expressed this way, we hope it is clear that value functions formalize a basic and familiar idea.

- Model of the environment is imitating of the behavior of the environment. Models are used for planning, by which we mean any way of deciding on a course of action by considering possible future situations before they are actually experienced. The incorporation of models and planning into reinforcement learning systems is a relatively new development. Modern reinforcement learning spans the spectrum from low-level ‘trial-and-error learning’ to high-level ‘deliberative planning’.

2.5.3 The Agent-Environment Interface

The learner who is the decision-maker is called the agent. Everything outside the agent that interact to the agent is called the environment where these interact continually and the agent selecting actions after that the environment responding to those actions and presenting new situations to the agent. The environment also gives rise to rewards, special numerical values that the agent tries to maximize over time. A complete specification of an environment defines a task, one instance of the reinforcement learning problem [19].

More specifically, the agent and environment interact at each of a sequence of discrete time steps as t = 1, 2, 3, 4 ……T, at each time step t, the agent receives some representation of the environment's state $s_t \in S$ where S is the set of possible states. After that, based on action algorithm the agent selects an action, $a_t \in A(s_t)$ where $A(s_t)$ is the set of actions available in state s_t. One time step later, in part as a consequence of its action, the agent receives a numerical reward $r_{t+1} \in R$ and finds itself in a new state , $r_{t+1.}$ Figure 2.16 diagrams the agent-environment interaction.

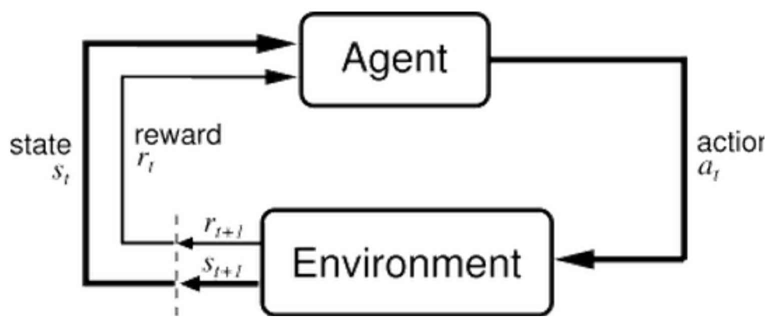


Figure 2.16: The agent-environment interaction in reinforcement learning [6]

At each time step, the agent implements a mapping from states to probabilities of selecting each possible action. This mapping is called the agent's policy and is denoted π_t where $\pi_t(s, a)$ is the probability that $a_t = a$ if $s_t = s$. Reinforcement learning methods specify how the agent changes its policy as a result of its experience. The agent's goal, roughly speaking, is to maximize the total amount of reward it receives over the long run [19].

The reinforcement learning framework is a considerable abstraction of the problem of goal-directed learning from interaction. It proposes that whatever the details of the sensory, memory, and control apparatus, and whatever objective one is trying to achieve, any problem of learning goal-directed behavior can be reduced to three signals passing back and forth between an agent and its environment: one signal to represent the choices made by the agent (the actions), one signal to represent the basis on which the choices are made (the states), and one signal to define the agent's goal (the rewards). This framework may not be sufficient to represent all decision learning problems usefully, but it has proved to be widely useful and applicable [19].

The environment in which the multi agent system is operating is dynamic due to the characteristics of the mobile wireless scenario e.g., existence of lognormal shadowing, fading, mobility and user, etc, and to the cross dependencies of actions made by the multiple agents. Thus, this scenario modeled by a stochastic game. A stochastic game is the extension of Markov Decision Processes (MDPs), which are the natural model of a single agent scenario, to multiple agents [20].

2.5.4 Goals and Rewards

The goal of an agent is to maximize the reward it receives in the long run which the reward signal passing from the environment to the agent. At each time step, the reward is a simple number, $r_t \in \mathbb{R}$ which total reward $t_T = r_t + r_{t+1} + r_{t+2} + r_{t+3} \ldots\ldots$. Informally, the agent's goal is to maximize the total amount of reward it receives. This means maximizing not immediate reward, but cumulative reward in the long run.

If we want agent to do something for us, we must provide rewards to it in such a way that in maximizing them the agent will also achieve our goals. Thus, the rewards we set up truly indicate what we want accomplished. In particular, the reward signal is not the place to impart to the agent prior knowledge about how to achieve what we want it to do. Therefore, if the agent achieved sub goals were rewarded, then the agent might didn’t achieve the real goal. Otherwise, the agent achieved the real goals may achieving sub goals. For example, a chess playing agent should be rewarded only for actually winning, not for achieving sub goals such taking its opponent's pieces or gaining control of the center of the board.

2.5.5 Learning in single-agent systems

A Markov Decision Processes MDP is a discrete time stochastic optimal control problem where it provides a mathematical framework for modeling decision-making processes in situations where outcomes are partly random and partly under the control of the decision maker. Here, operators take the form of actions, i.e. inputs to a dynamic system, which probabilistically determine successor states. A MDP is defined in terms of a discrete-time stochastic dynamic system with finite state set $S = \{s_1, \ldots, s_k\}$. As we mentioned in section 2.5.3 time is represented by a sequence of time steps, t = 0, 1, . . . ,T. At each time step, the agent or controller observes the system's current state and selects an action, which is executed by being applied as input to the system when the controller executes action a_t , the system state at the next step changes from s to v, with a state transition probability $P_{s,v}$. We further assume that the application of action 'a' in state's' incurs an immediate cost c(s, a). When necessary, we refer to states, actions, and immediate costs by the time steps at which they occur, by using s_t, at and c_t, where $a_t \in A$, $s_t \in S$ and $c_t = c(s_t, a_t)$ are, respectively, the state, action and cost at time step t. To sum up, a MDP consists of [19]:

- A set of states, S.
- A set of actions, A.
- A cost function C: $S \times A \in R$.
- A state transition function P: $S \times A$ □$\Pi(S)$, where a member of $\Pi(S)$ is a probability distribution over the set, S.

Markov model is the model if the state transitions are independent of any previous environment states or agent actions.

RL problems model the world using MDP formulism. In the literature, three ways have been identified to solve RL problems. The first one consists of the knowledge of the state transition probability function from state 's' to state 'v', Ps,v(a), and is based on dynamic programming. The second and third forms to solve RL problems. As a result, Monte Carlo (Monte-Carlo search framework (Tesauro and Galperin, 1996), which has been successfully applied to complex computer games such as Go, Poker, Scrabble, multi-player card games) and TD are primarily concerned with how an agent ought to take actions in an environment so as to minimize the notion of long-term cost or maximization return, that is, so as to obtain the optimal policy. When state transition probability is not known, but a sample transition model of states, actions and costs can be built, Monte Carlo methods can be applied to solve the MDP problem.

2.5.6 Optimal Value Functions and TD Policy

TD methods combine elements of dynamic programming DP and Monte Carlo ideas, they learn directly from experience which is a characteristic of Monte Carlo methods and they gradually update prior estimate values, which is common of dynamic programming. TD methods allow an online learning which is crucial for long-term/continues applications.

RL algorithms are based on the computation of value functions “the state-value function” V(s), or the state-action value function, Q(s, a), which measure how good is for an agent to be in a given state or to execute an action in a given state if based on the future expected cost. Thus, the expected costs for the agent in the future are given by the actions it will take and therefore, the value functions depend on the policies being followed. The state-value of state‘s’ is defined as the expected infinite discounted sum of costs that the agent gains if it starts in state ‘s’ and then executes the complete decision policy π,

$$V^{\pi}(s) = E_{\pi}\left\{\sum_{t=0}^{\infty} \gamma^{t} c_{t} \mid s_{t} = s\right\} \qquad 2.8$$

where $0 < \gamma < 1$ is a discount factor which determines how much expected future costs affect decisions made now. Similarly, the Q-value $Q(s,a)$ represents the expected decreased cost for executing action ‘a’ at state s and then following policy π thereafter.

$$Q^{\pi}(s,a) = E_{\pi}\left\{\left\{\sum_{t=0}^{\infty} \gamma^{t} c_{t} \middle| s_{t} = s, a_{t} = a\right\}\right\} \qquad 2.9$$

There is at least one policy that is better than or equal to all other policies, i.e., an optimal policy, π^{*}. The optimal policy shares the optimal state value function V^{*} and the optimal state-action value function Q^{*} and can express as:

$V^{*}(s) = \max_{\pi} V^{\pi}(s)$ and $Q^{*} = max_{\pi} Q^{\pi}(s,a)$

The optimal value function can be expressed recursively with the Bellman optimality equations as [21]:

$$V^{*}(s) = max_{a} \sum_{v} P(v \mid s,a)[R(v|s,a) + \gamma V^{*}(v)] \qquad 2.10$$

Similarly, for Q values: $Q^{*}(s,a) = \sum_{v} P(v \mid s,a)[R(v|s,a) + \gamma V^{*}(v)]$

The policy maps states to actions, hence, it determines the actions the agent will follow depending on the given situation. Policies can be classified into two groups.

1) The behavior policy, which determines the conduct of the agent, i.e. the actual action selected by the agent in the current state.

2) The estimation policy, which determines the policy evaluated, or the action in the next state used for the evaluation of behavior policy. In RL there are two methods to ensure a sufficient exploration, the on-policy and off-policy methods, which basically differ on the form they select the estimation policy.

On-policy methods: These methods evaluate or improve the policy, π, they estimate the value of a policy while using it for control. This means that, the policy followed by the agent to select its behavior in a given state (behavior policy) is the same used to select the action (estimation policy) based on which it evaluates the behavior followed.

• Off-policy methods: These methods do distinguish between behavior and estimation policies. Therefore, the policy to generate behavior, π is unrelated to the policy evaluated [6]. The policy evaluated in off-policy methods is the one corresponding to the best action in the next state, π*, given the current agent experience. The correct selection of these parameters highly influences the performance of the learning process.

2.5.7 Temporal Difference Learning

Both TD and Monte Carlo methods use experience to solve the prediction problem. Given some experience following a policy, both methods update their estimate V of V^{π}. If a nonterminal state s_t is visited at time t, then both methods update their estimate based on what happens after that visit. Roughly speaking, Monte Carlo methods wait until the return following the visit is known, then use that return as a target for $V(s_t)$. A simplest method $TD(0)$ update $V(s_t)$ is 2.11

$$V(s_t) \leftarrow V(s_t) + \alpha[r_t + \gamma V(s_{t+1}) - V(s_t)]$$

where α is the "learning rate" and γ is the "discount rate' are constant ranged from 0 to 1. In TD learning, agents attempt to select actions that minimize the discounted costs they receive over the future, which is why the discount rate, γ. TD methods need wait only until the next time step. At time t+1 they immediately form a target and make a useful update using the observed reward r_{t+1} and estimate $V(s_{t+1})$[19].

2.5.8 Exploration and Exploitation

Exploration the environment is one important aspect in TL where it important to gather information in order to build a policy. We do not want leave unexplored areas but we also want to use the accumulated knowledge to make better decisions. To gain more rewards the agent must balances between exploration and exploitation. In many cases the exploration strategy depeneds on the time that the agent has interested with environment, some common strategies to select actions and explore environment as follow [21]

- $\acute{\varepsilon}$ -greedy: where most of the time the selected action is the one with the largest estimated accumulated reward but with probability $\acute{\varepsilon}$ an action randomly selected.
- Softmax: where the probability of selecting an action depends on its estimated accumulated reward. The most common being the Boltzmann or Gibbs distribution, which selects an action on state s with probability:

$$\frac{e^{Q(s,a)/\tau}}{\sum_{b=1}^{n} e^{Q(s,b)/\tau}} \qquad 2.12$$

where τ is a positive number (temperature) and n is a number of actions [3].

2.5.9 TD Q-learning method

In QL, the state of the surrounding environment and available actions to the agents are commonly represented by discrete sets. The QL algorithm is based on quantifying, by means of the Q-function, the quality of an action in a certain state. Therefore, to be able to learn from the past, the Q-values have to be stored in a representation mechanism. The lookup table, represented in Figure 2.17, is the most commonly used and the most direct method when the memory requirement is not a problem. The Q-value function is

$$Q(s,a) \Rightarrow Q(s,a) + \alpha(r + \gamma\max\big(Q(v,a)\big) - Q(s,a)) \qquad 2.13$$

where $\max\big(Q(v,a)\big)$ is the maximum Q-value of state v which state transition from s to v if did action, a [21].

s_1	$Q(s_1,a_1)$		$Q(s_1,a_l)$
s_2	$Q(s_2,a_1)$		$Q(s_2,a_l)$
$\vdots$			
s_k	$Q(s_k,a_1)$		$Q(s_k,a_l)$

Figure-2.17: Q-learning Table [22].

Bibliography

[1] Abd El-Gawad , M. Tantawy and ?. El-Mahallawy," LTE QOS Dynamic Resource Block Allocation with power source limitation and queue stability constrains," International Journal of Computer Networks & Communications (IJCNC), Vol.5, No.3, May 2013.

[2] Jie Zhang, Guillaume de la Roche," Femtocells: Technologies and Deployment ", 2010 John Wiley & Sons Ltd, ISBN: 978-0-470-74298-3

[3] Eduardo F. Morales, "Introduction Reinforcement Learning", National Institute of Astrophysics, Optics and Electronics, México,2012.

[4] Ramjee Prasad, Fernando J. Vele," WiMAX Networks: Techno-Economic Vision and Challenges,",Springer Science & Business Media, Jun 10, 2010 - Technology & Engineering - 516 pages.

[5] Ramjee Prasad,' OFDM for Wireless Communications Systems,' Artech House universal personal communications series Includes bibliographical references and index.ISBN 1-58053-796-0.

[6] Weinstein,and S., & Ebert "Data transmission by frequency-division multiplexing using the discrete Fourier transform," Communication Technology, IEEE, P. (1971).

[7] T. Deepa, and R. Kumar," An Efficient Layered FFT Approach for Reduction of PAPR, Spectral Re-growth and CFO on the Performance of an OFDM System" , Wseas Transactions on Communication, E-ISSN: 2224-2864, 2014 .

[8]" LTE Overview " [Online]. Available: http://www.3gpp.org/technologies/keywords-acronyms/98-lte ,2014.

[9] Harri Holma and Antti Toskala ,"LTE for UMTS: OFDMA and SC-FDMA Based Radio Access", 2009 John Wiley & Sons, Ltd. ISBN: 978-0-470-99401-6.

[10] Hamza, J., Long Term Evolution (LTE) - A Tutorial, Network System Labora- tory, Simon Fraser University, Canada, October 13, 2009.

[11] A. Z. Yonis," Downlink and Uplink Physical Channels in Long Term Evolution", I.J. Information Technology and Computer Science, 2012, 11, 1-10.

[12] S.Hussain," Dynamic Radio Resource Management in 3GPP LTE", This thesis is presented as part of Degree of Master of Science in Electrical Engineering, Thesis Number: MEE09:06, 2009.

[13] Juan J. Sánchez, D. Morales-Jiménez, G. Gómez, J. T,” Physical Layer Performance of Long Term Evolution Cellular Technology”, Supported by the Spanish Government and the European Union under project TIC2003- 07819 (FEDER) and by the company AT4W (former CETECOM, S.A.), 2007.

[14] Jie Zhang, Guillaume de la Roche,” Femtocells: Technologies and Deployment “, Wiley & Sons Ltd, ISBN: 978-0-470-74298-3 , 2010.

[15] Mahmoud Ouda,” Interference-Optimal Frequency Allocation in Femtocellular Networks”, A thesis submitted University Kingston, Ontario, Canada March 2012.

[16] N.Saquib, E. Hossain, L. Baole and D. Kim,"Interference Management of OFDMA in Femtocell Networks: Issues and Approach", IEEE Wireless Communications-Journal. June 2012.

[17] Andrea Goldsmith,” Wireless Communication”, Stanford University.

[18] Theodore Rappaport,” Wireless Communications: Principles and Practice”, Prentice Hall PTR Upper Saddle River, NJ, USA ©2001 ISBN:0130422320.

[19] Richard S. Sutton and Andrew G. Barto ,”Reinforcement Learning: An Introduction” ,The MIT Press. Cambridge, Massachusetts London, England.

[20] G. Chalkiadakis, “Multiagent reinforcement learning: Stochastic games with multiple learning players,” University of Toronto, Tech. Rep., 2003.

[21] Ana Galindo-Serrano and Lorenza Giupponi,” Femtocell Systems with Self Organization Capabilities”, Centre Tecnologic de Telecomunicacions de Catalunya (CTTC) Parc Mediterrani de la Tecnologia, Av. Carl Friedrich Gauss 7, Barcelona, Spain , 2011.

Chapter 3

Proposed Scheme

This chapter describes the problem context in detail, in terms of technologies and infrastructure assumed. It also discusses how resources are assigned. Moreover, it describes the nature of the self-organization problem and modeling it to an assignment problem which this is proposed by reinforcement learning (Q-learning) to solve resource allocation problem.

3.1.1 System Description

This study is concerned with downlink interference mitigation by proper frequency allocation in OFDMA networks. Mixed OFDMA networks comprising femto-tiers layering macro-tiers are considered. The preceding statements can be broken down into a set of characteristics that describes the environment upon which the proposed scheme should be run:

- OFDMA System. As the air interface technology of choice for downlink in LTE, OFDMA is considered in simulation.
- Downlink. The study focuses on downlink direction instead of both, downlink and uplink directions. Specialty in this matter allowed for the development of a tailored powerful algorithm.
- Deployment Model. Practicality instructions that the proposed learning algorithm supports tiered deployment, where the scenario of the models are enterprise femtocells and femtocells underlay macrocells networks.

High Level Operation

In an OFDMA system with mixed BSs - (HeNBs and MeNBs) – the bandwidth is shared between the HeNB and MeNB layers. Each BS has a finite number of Resource Blocks (RBs) representing the available spectrum, and has a number of UEs uniformly distributed within the cell radius. At each BS, downlink SINR is calculated. For each femtocell, the Q-table created for each user as shown in figure 2.17 where each raw represents by Q-value for all actions of specific state. A centralized entity aggregates all. ϵ-greedy algorithm is then run to produce an optimal solution in terms of maximizing the total capacity.

3.1.2 Simulation Scenario

Consider the aforementioned OFDMA network with two layers, a HeNB layer within a MeNB layer but we proposed two scenarios for employment HeNB.

1. First one is enterprise femtocell network as shown in figure 3.1. Thus, in this scenario there is not any interference effect of MeNB to femto-users and also there is not any interference of HeNB to macro-users "cross-tier interference" vice versa. Therefore, in this scenario the interference effective is from each HeNB to other femto-users "co-tier interference". In this model, the simulated office building contains open planned offices in partitioned cubicles where its size is 50 m×50 m and divided to 8 blocks apartments. We assume that 8 HeNBs are deployed in the building with some rudimentary planning one HeNB for each section, but without performing a detailed cell planning survey of the building. Since the distribution of users throughout the building is relatively even, the femtocells are deployed roughly equally apart.
2. The second scenario is HeNBs within MeNB. Whereas the femtocellular layer consists of several shorter range cells resulting from the deployment of HeNBs in an ad-hoc manner. Therefore, we consider in this model 1 MeNB with radius D = 400 m is underlaid with 2 HeNBs as shown in figure 3.2 where HeNBs implemented in two rooms where them size is 20m x 10m. In this state, HeNB share part of resource block that used in MeNB. Thus, two types of interference are appeared, co-tier and cross-tier interferences.

Each eNB has a number of users (UEs) to serve. UEs that belong to the femtocellular layer, i.e., connected to HeNBs, are referred to as Femtocell User Equipments (FUEs), whereas UEs connected to a MeNB are referred to as Macrocell User Equipments (MUEs).

As we mentioned in section 2.1.1, UEs are served in units of RBs. In OFDMA downlink, a RB is a basic time-frequency unit (see Figure 2.10), it consists of 12 consecutive subcarriers in the frequency domain and 7 OFDM symbols in the time domain, assuming normal cyclic prefix. The minimum unit to serve a UE is a RB, that is, a user is either admitted access to a whole RB or not. The problem investigated in this study is HeNB self-configuration to assignment of resource blocks among system users. An optimal assignment algorithm is proposed, aimed at maximizing the system capacity in two scenarios.

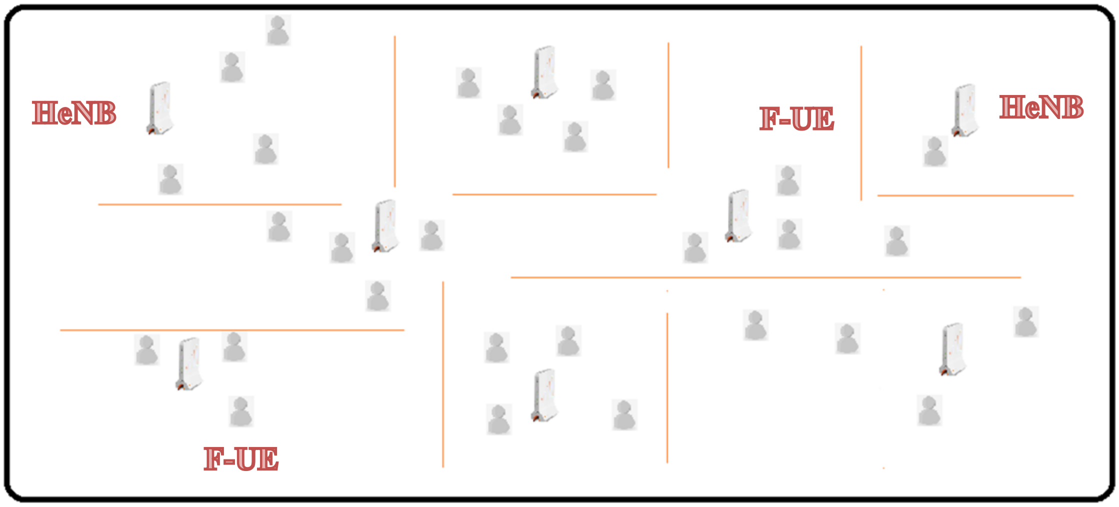


Figure 3.1: Enterprise Femtocell Network.

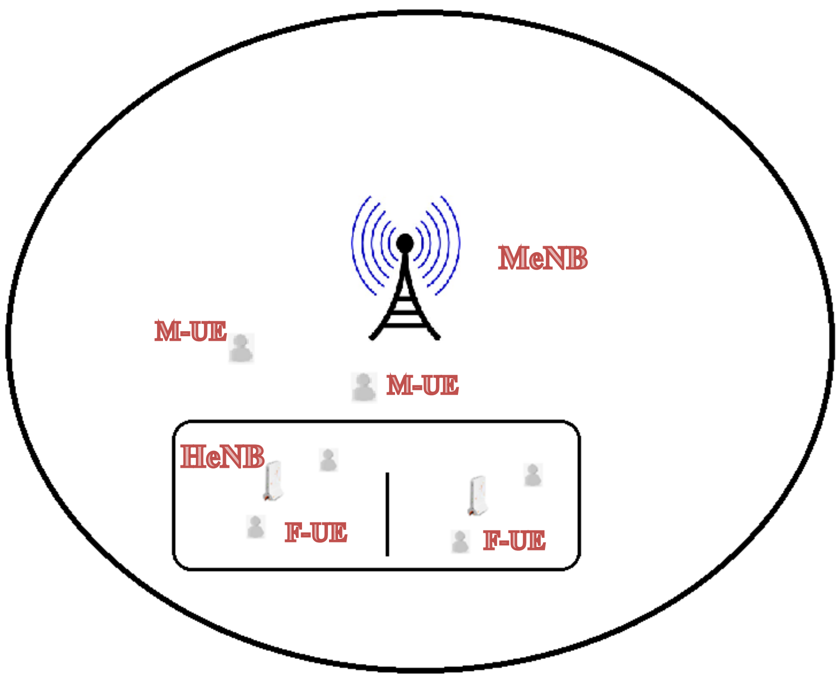


Figure 3.2: HeNB underlying MeNB network.

3.2.1 System Model

We assume that both MeNB and HeNBs operate in the same frequency band and have the same amount R of available RBs. The work presented in this thesis focuses only on the downlink operation. We denote by $P^{f,F} = (p_1^{f,F}, \dots\dots, p_R^{f,F})$ and $P^{m,M} = (p_1^{m,M}, \dots\dots, p_R^{m,M})$ the transmission power vector of HeNB f and MeNB m with $p_r^{f,F}$ and $p_r^{m,M}$ denoting the downlink transmission power of HeNB and MeNB in RB r, respectively. The maximum transmission power for HeNB is P_{max}^{F} which the total power assigned for RBs, femto-users, should be less than maximum power $\sum_{r=1}^{R} p_r^{f,F} \leq P_{max}^{F}$.

The SINR at MUE $u^m \in U_m$ allocated in RB r of MeNB m is

$$SINR_r^m = \frac{p_r^{m,M} h_{mm,r}^{MM}}{\sum_{k=1,k\neq m}^{M} p_r^{k,M} h_{km,r}^{MM} + \sum_{f=1}^{N} p_r^{f,F} h_{fm,r}^{FM} + \sigma^2} \quad 3.1$$

with $m = 1,\dots,$ M. here, $h_{mm,r}^{MM}$ indicates the link gain between the transmitting MeNB m and its user u^m , $h_{km,r}^{MM}$ indicates the link gain between the transmitting MeNB k and MUS u^m in MeNB m and $h_{fm,r}^{FM}$ indicates the link gain between the transmitting HeNB f and MUS u^m of MeNB m. Finally, σ^2 is the noise power. The capacity of MeNB m is:

$$c^{m,M} = \sum_{r=1}^{R} \frac{BW}{R} \log_2(1 + SINR_r^M) \quad , \quad 3.2$$

with $m = 1, \dots, M$.

The SINR at FUS $u^f \in U_f$ allocated in RB r of HeNB f is:

$$SINR_r^f = \frac{p_r^{f,F} h_{ff,r}^{FF}}{\sum_{m=1}^{M} p_r^{m,M} h_{mf,r}^{MF} + \sum_{k=1,k\neq f}^{N} p_r^{k,F} h_{kf,r}^{FF} + \sigma^2} \quad 3.3$$

With f= 1, ..., N. here, $h_{ff,r}^{FF}$ denotes the link gain between the transmitting HeNB f and its FUS u^f, $h_{mf,r}^{MF}$ indicates the link gain between the MeNB m and FUS u^f in HeNB f and $h_{kf,r}^{FF}$ denotes the link gain between the transmitting HeNB k and UE u^f of HeNB f.

The capacity of HeNB f is:

$$c^{f,F} = \sum_{r=1}^{R} \frac{BW}{R} \log_2(1 + SINR_r^f) \quad 3.4$$

With $f = 1, \dots \dots, N$

3.2.2 Simulation Parameters

MeNB and HeNB systems are operating at 1850 MHz and to be based on LTE. Therefore, the frequency band is divided into RBs. For simulations, we consider an amount of RBs R = 15 for scenario-1 and R=6 for scenario-2 as talk about in section-3.1.2, which corresponds to the minimum LTE implementation and corresponds to a channel bandwidth of BW = 3 MHz and 1.6 MHz, respectively. The antenna patterns for eNB, HeNB and eNB/HeNB UEs are omnidirectional, with 18 dB_i, 0 dB_i and 0 dB_i antenna gains, respectively. The shadowing standard deviation is 8 dB and 4 dB, for MeNB and HeNB systems, respectively. The MeNB and HeNB noise figures are 5 dB and 8 dB, respectively. The transmission power of the MeNB is 40 dB_m, whereas the HeNB adjusts its power and allocates RBs through the learning scheme where a total maximum value of P_{max}^F =15 dB_m. The considered PL models are for two scenarios and are summarized in Table 3.1[27].. WP_{out} = 15 dB and WP_{in} = 5 dB are the penetration losses of the building.

MeNB to HUEs	Indoor	$PL(dB) = 15.3 + 37.6 \log_{10} d + WP_{out}$
HeNB to HUEs	Indoor	$PL(dB) = 38.46 + 20 \log_{10} d + .7 d_{indoor} + w_p WP_{in}$
HeNB to MUEs	Indoor	$PL(dB) = \max(15.3 + 37.6 \log_{10} d, 38.46 + 20 \log_{10} d) + .7 d_{indoor} + w_p WP_{in} + 2WP_{out}$

Table 3.1: Path Loss Models for Urban Deployment [27]

3.2.3 Q-learning for Resources Allocation Management

In this section we first define the states, actions and cost function, introduced in Section 3.1.1, which define the proposed learning algorithm. We present two case studies, the state and cost, used in case study 1 that applied for scenario-1, have been designed in order to highlight the ability of the learning to manages the resources allocation for each HeNB. Thus, the objective of this case is approaching to converge to desired SINR values where there is not connecting or transferring any data between any HeNBs. So, each HeNB is autonomous and has self responsibility to find the optimal resources allocation for own users. On the other hand, case study 2 that functional for scenario-2. Therefore, MeNB worked as static system for serving MUEs so in this learning case, applying in HeNBs, selected state representation and cost function with the purpose of highlight the ability of the learning to manage the resources allocation for each HeNB. Thus, the objective of this case is approaching to converge to desired SINR values in the macrocell and femtocell systems. But in this case there is transferring SINR of suffering MUEs to HeNBs. So, each HeNB is autonomous and has self responsibility to find the optimal resources allocation for own users with decreasing cross-tier interference.

1) Learning design: case study 1:

In our system the multiple agents with learning capabilities are the femtocell BSs, so that for each HUE, the agents are in charge of identifying the current environment state, select the action based on the action selection policy and execute it. In the following, for each agent $f = 1, 2, \ldots, N$ and RBs $r = 1, 2, \ldots, R$ we define system state, action, associated cost and next state.

- **State:** The system state for agent f and $RB\ r$ is defined as:

$$s_{UE}^{f} = \{I_{UE}^{f}, Pow^{f}\}$$

3.5

Notice that, s_{UE}^{f} is fiting into a Q-table with k dimension where k is number of actions.

- **User Capacity Indicator:** $I_{UE}^{f} \epsilon I$ represents a binary indicator to specify whether the femtocell system is generating aggregated interference above or below the FUE required capacity threshold. This measure is based on the SINR value computed at the FUE allocated at $RB\ r$. The set of possible values is based on:

$$I_{UE}^{f} = \begin{cases} 0 & If\ c_{UE}^{f} > C_{thr} \\ 1 & otherwise \end{cases} \quad 3.6$$

Where c_{UE}^{f} is the capacity of femtouser and C_{thr} represents the minimum capacity value that can be received by FUE.

- **Femtocell total transmission power indicator:** One of the requirements of femtocells is to transmit at low-power levels, for this reason we include in the state definition the Pow^{f} indicator, in order to guarantee that the femtocell total transmission power over all RBs, Pow^{f} , is below the threshold P_{max}^{F}. It is given by:

$$Pow^{f} = \sum_{r=1}^{R} p_{r}^{f} \quad 3.7$$

As a result, Pow^{f} is a binary indicatror defined as follow:

$$Pow^{f} = \begin{cases} 0 & if\ Pow^{f} < Pow_{max}^{F} \\ 1 & otherwise \end{cases} \quad 3.8$$

Therefore, in this case study the state of the environment is characterized by k = 4 possible situations.

- **Actions:** The set of possible actions are the l power levels and R RB A = {(p_1, r_1) , (p_1, r_2) , … , $(p_{1,} r_R)$, (p_2, r_1) , … , (p_2, r_R) , … , (p_l, r_1) , … , (p_l, r_R) } that femtocell can assign to femtousers FUE.

- **Cost:** The cost c assesses the immediate return incurred due to the assignment of action a at state s. The considered cost function is:

$$c = \begin{cases} 90 + SINR_{UE} & if\ s_{UE}^{f} = \{I_{UE}^{f} = 0, Pow^{f} = 0\} \\ -10 & otherwise \end{cases} \quad 3.9$$

The rationale behind this cost function is that the Q-learning aims maximize the capacity with keep the total power under power threshold that done by the mean part of Q-learning ,cost, here in state $s_{UE}^{f}(0,0)$ the cost function increases respect to SINR when there is interference that reduces lower cost than the state without interference. Otherwise, the cost function equal -10 that mean this state-action is not efficient because whether total power more than threshold or the user capacity lower than threshold.

- Next State: The state transition from s to v in RB FUE is determined by the learning-based resources allocation (power and RB).

2) Learning design: case study 2:

Differently from case study 1, here we suppose different cost definition and maintain the state, action and next state as in study 1 since we furthermore consider a macrocell system performance. We define the cost respect macro user SINR as follows.

Cost: The cost c assesses the immediate return incurred due to the assignment of a certain action in a given state. The considered cost function is:

$$c = \begin{cases} 90 + SINR_{UE} & if\ s_{UE}^{f} = \{I_{UE}^{f} = 0, Pow^{f} = 0\} \\ -20 & if\ c_{UE}^{m} < C_{thr} \\ -10 & otherwise \end{cases} \quad 3.10$$

The rationale behind this cost function, as we mentioned in case 1, is that the Q-learning aims maximize the capacity for femtouser additionally in this case maximize the capacity for macrouser c_{UE}^{m} with keep the total power under power threshold.

In order to make clearer how states are quantized to a k-dimensional space for agents to represent the knowledge they acquire on the run, table 3.2 represents the Q-table for case study 2.

$s_{1=}\begin{Bmatrix} I_{UE}^{f}=0 \\ Pow^{f}=0 \end{Bmatrix}, Q(s_1, a_{1,1})$	......	$Q(s_{1,}a_{l,R})$
$s_2 = \begin{Bmatrix} I_{UE}^{f}=0 \\ Pow^{f}=1 \end{Bmatrix}, Q(s_2, a_{1,1})$	......	$Q(s_2, a_{l,R})$
$s_3 = \begin{Bmatrix} I_{UE}^{f}=1 \\ Pow^{f}=0 \end{Bmatrix}, Q(s_3, a_{1,1})$	......	$Q(s_{3,}a_{l,R})$
$s_4 = \begin{Bmatrix} I_{UE}^{f}=1 \\ Pow^{f}=1 \end{Bmatrix}, Q(s_4, a_{1,1})$	......	$Q(s_{4,}a_{l,R})$

Table 3.2: Q-table for task UE of agent f, for case study 2

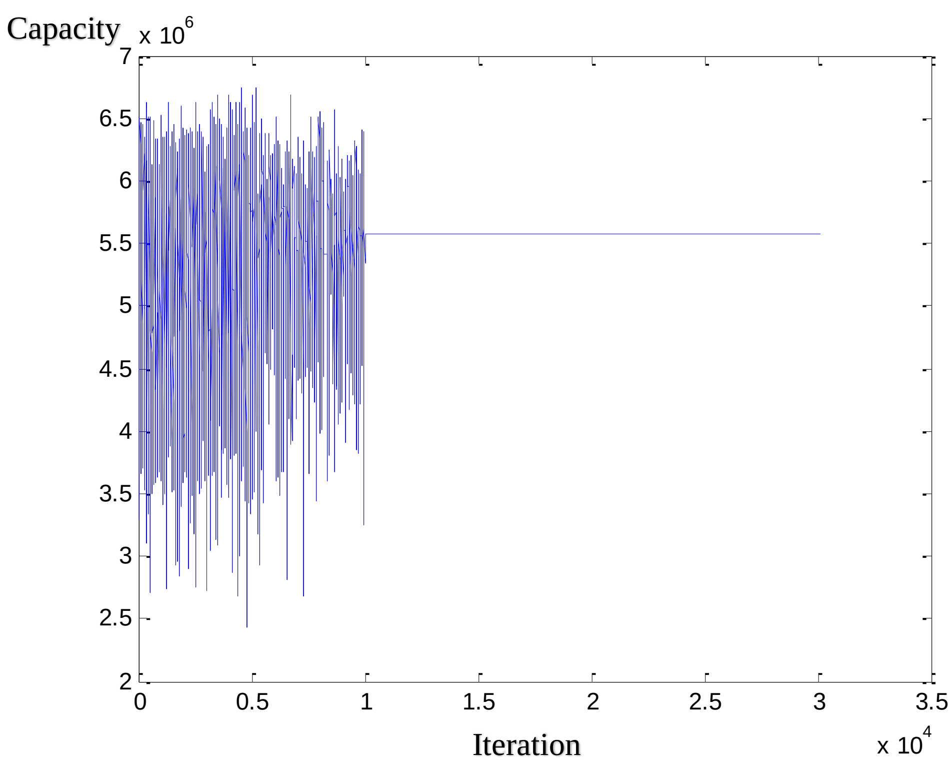
Finally, mentioning that we assume that the c_{UE}^{f}, $SINR_{UE}$ and c_{UE}^{m} can be computed by the macrocell and the femtocell, respectively, based on the Reference Signal Received Quality (RSRQ) reported by the UEs allocated in RB r. The RSRQ is the quantification of the user received signal considering both, signal strength and interference [10].

3.3 Simulation Results

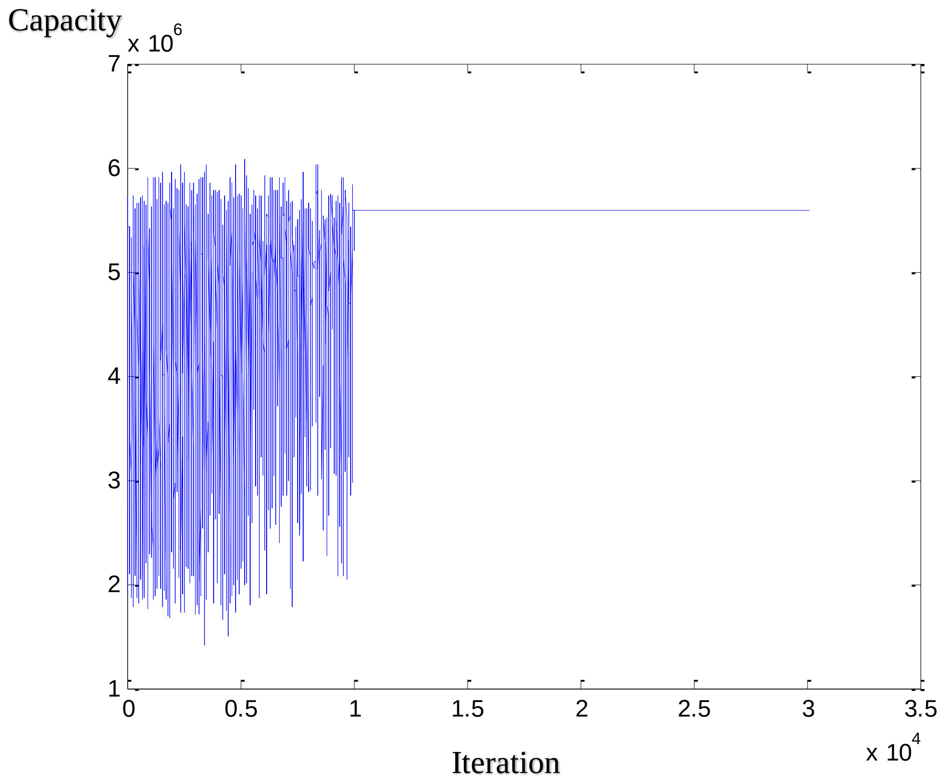
3.3.1 Scenario-1 Results

Results presented in this section correspond to a learning process modeled following the case study 1 and with an ε-greedy constant action selection policy. In that order, ε-greedy algorithm is responsible of exploring and exploiting optimal action. In this state, we supposed ε value in three stages (.2, .005 and .000005) where we generated random number if it less than ε-greedy the action algorithm exploration otherwise exploitation.

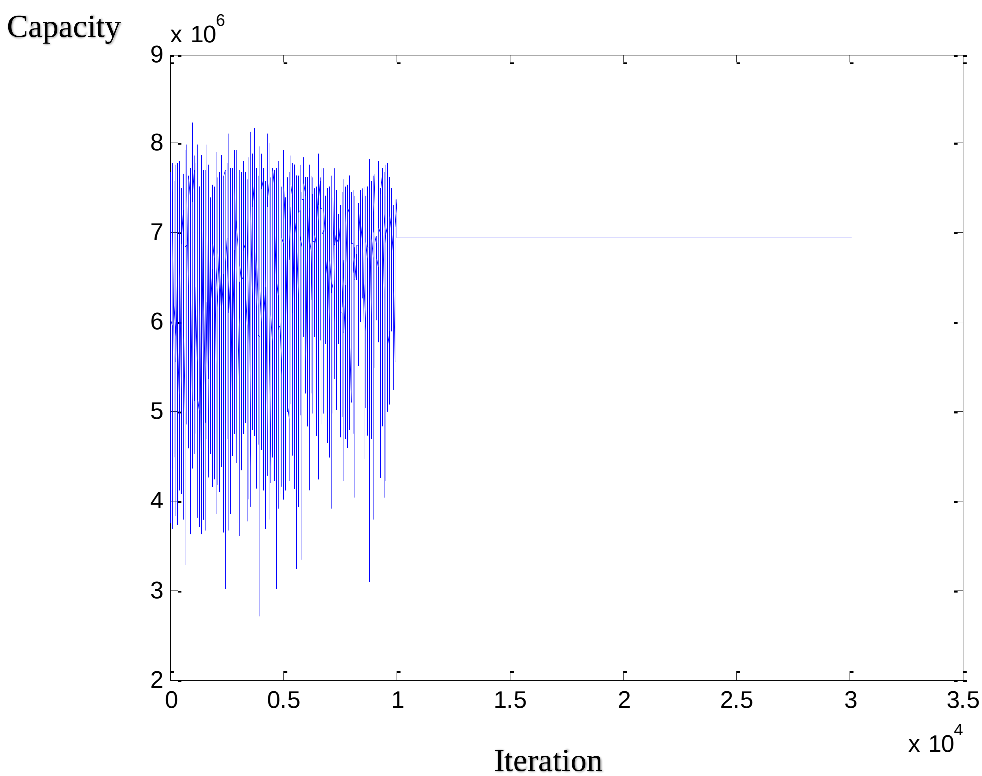
First of all, it has to be noted that the decentralized Q-learning algorithm, as any other learning scheme, needs a learning phase to learn the optimal decision policies. However, once completed the learning process and acquired the optimal policy. Figure 3.3 show the capacity of all 8 HeNBs as a function of the learning iterations. They can be observed that each HeNB finds optimal resource allocation for own users. So, the optimal actions ensure the efficient serves for users and do not exceed on the other users registered on different HeNB. Moreover, figure 3.4 show the power of the HeNBs as a function of learning iterations. Therefore, they can be observed, as we mentioned after each HeNB find optimal actions, the total power estimated to serve user without exceeding to maximum power threshold.



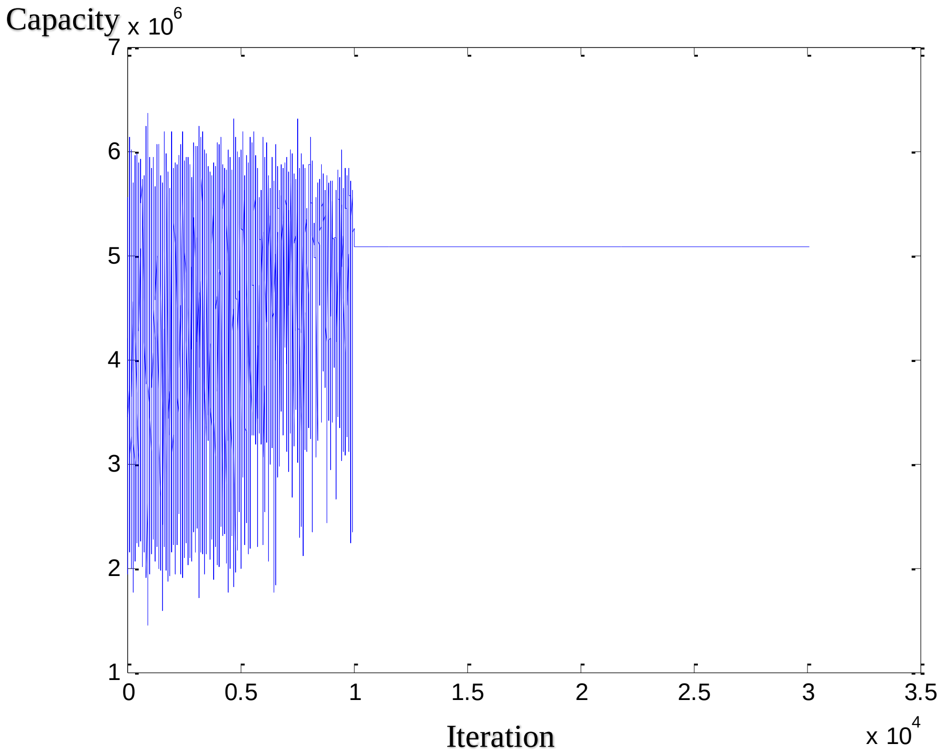
(1)


Capacity
x 10^6
7
6
5
4
3
2
1
0
0.5
1
1.5
2
2.5
3
3.5
x 10^4
Iteration

(2)

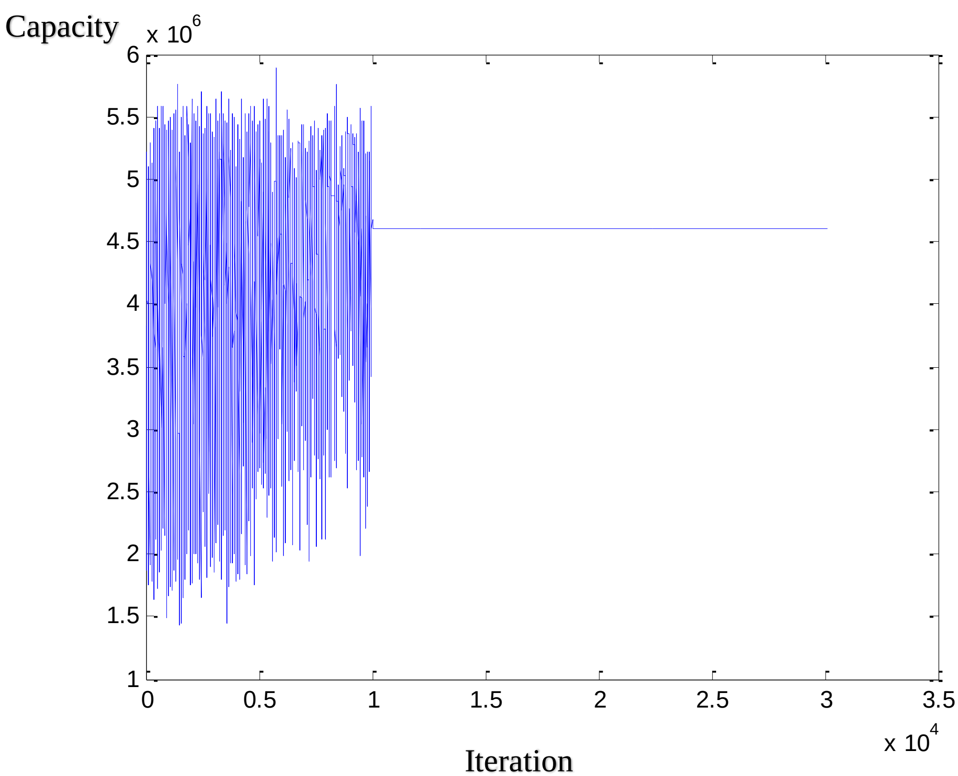

Capacity
x 10^6
9
8
7
6
5
4
3
2
0
0.5
1
1.5
2
2.5
3
3.5
x 10^4
Iteration

(3)



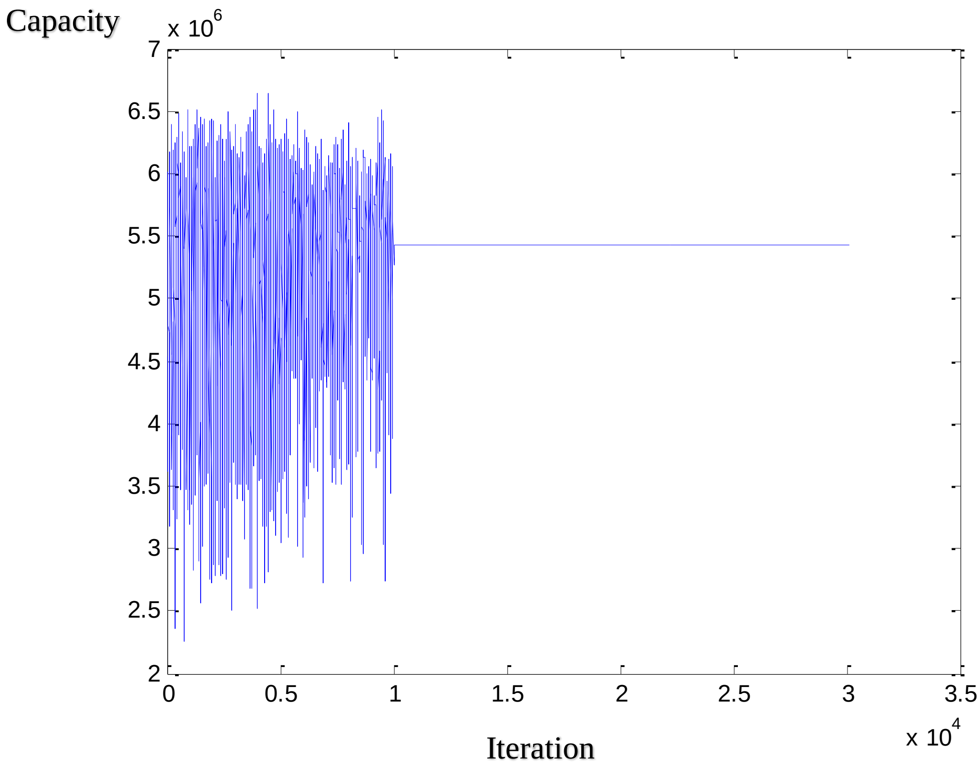
Capacity
x 10^6
7
6
5
4
3
2
1
0
0.5
1
1.5
2
2.5
3
3.5
Iteration
x 10^4

(4)

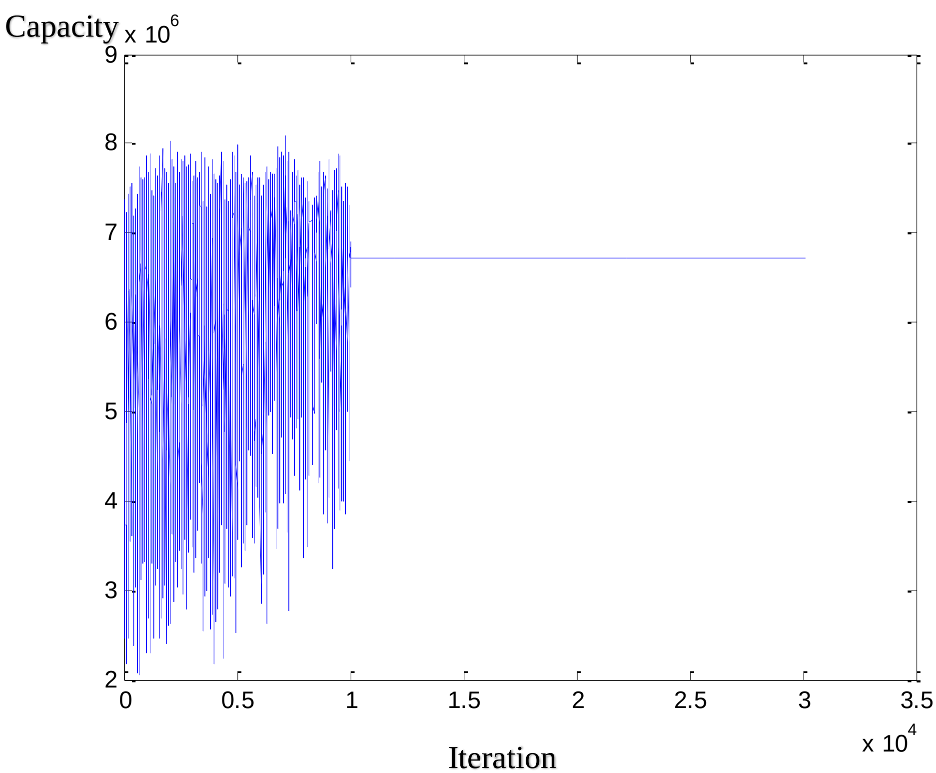


Capacity
x 10^6
6
5.5
5
4.5
4
3.5
3
2.5
2
1.5
1
0
0.5
1
1.5
2
2.5
3
3.5
x 10^4
Iteration

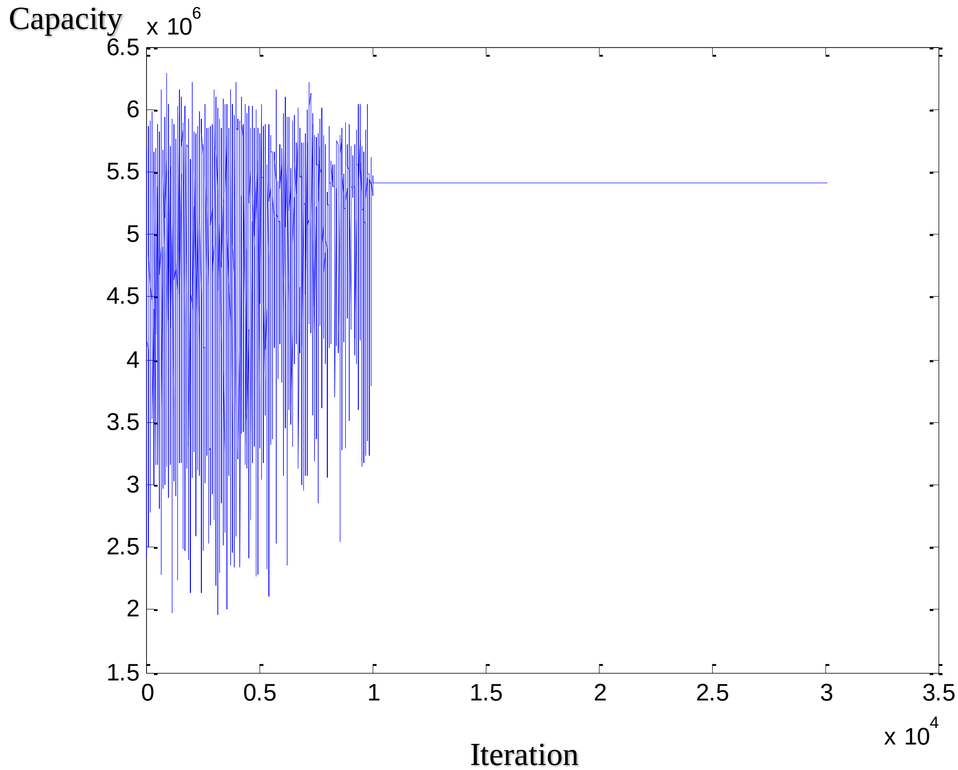
(5)


Capacity
x 10^6
7
6.5
6
5.5
5
4.5
4
3.5
3
2.5
2
0
0.5
1
1.5
2
2.5
3
3.5
x 10^4
Iteration

(6)

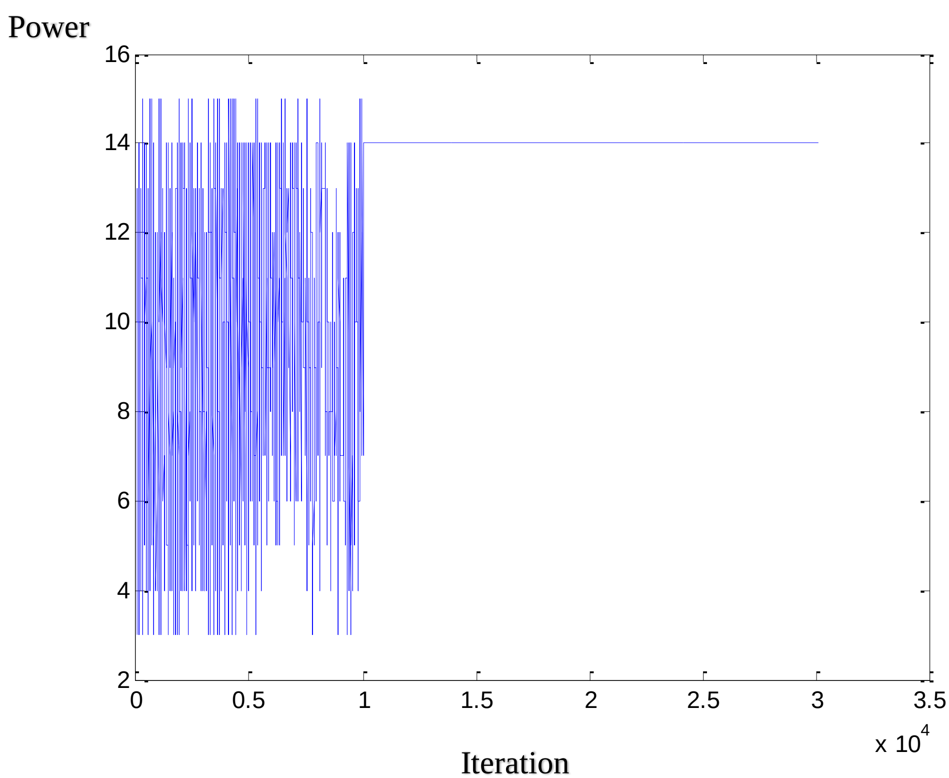

Capacity
x 10^6
9
8
7
6
5
4
3
2
0
0.5
1
1.5
2
2.5
3
3.5
x 10^4
Iteration

(7)

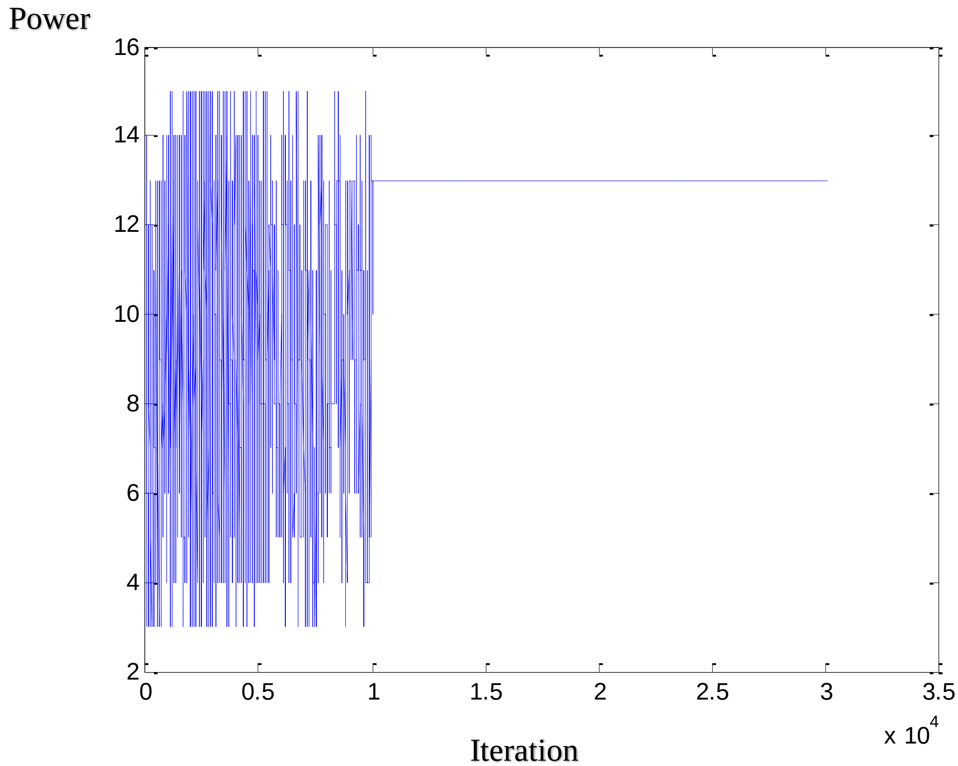


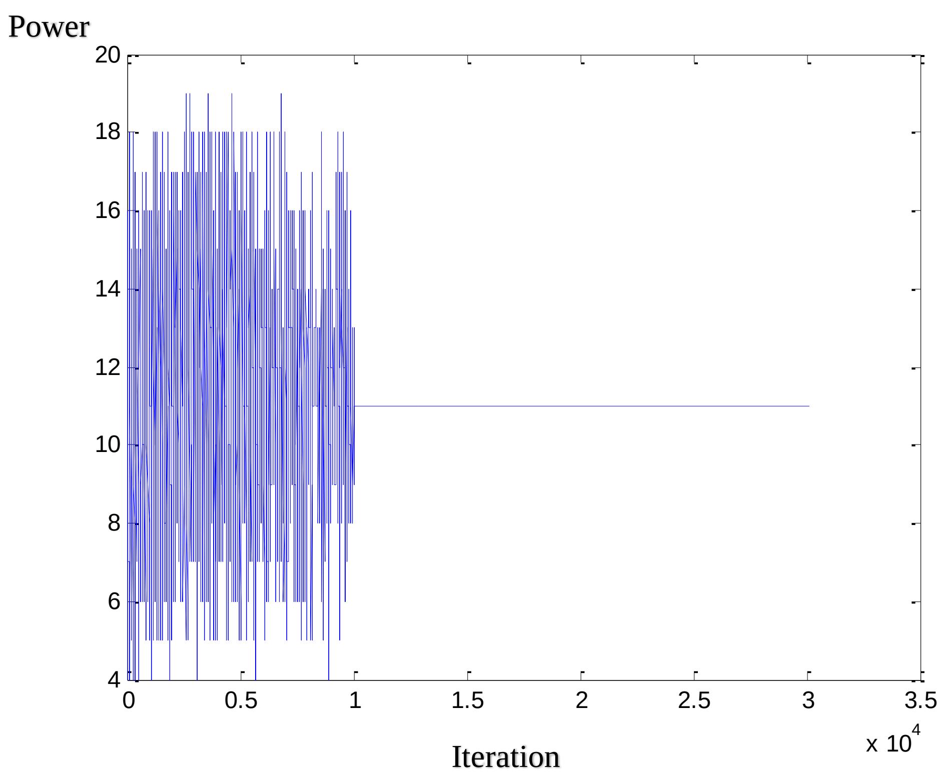
(8)

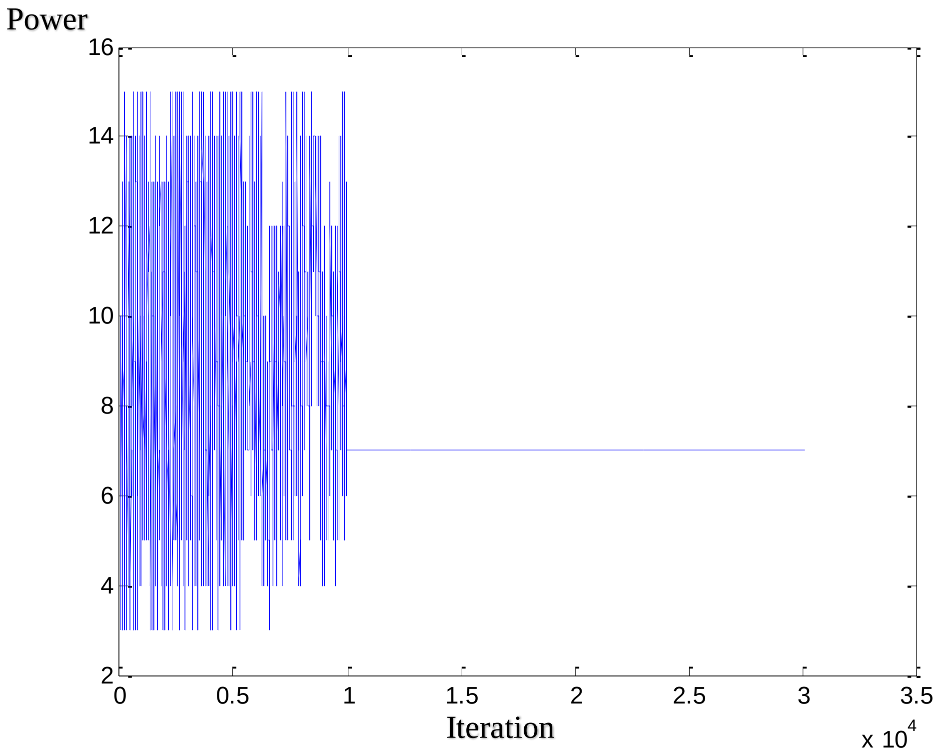
Figure 3.3: Femtocell Capacity as Function of Learning Iterations



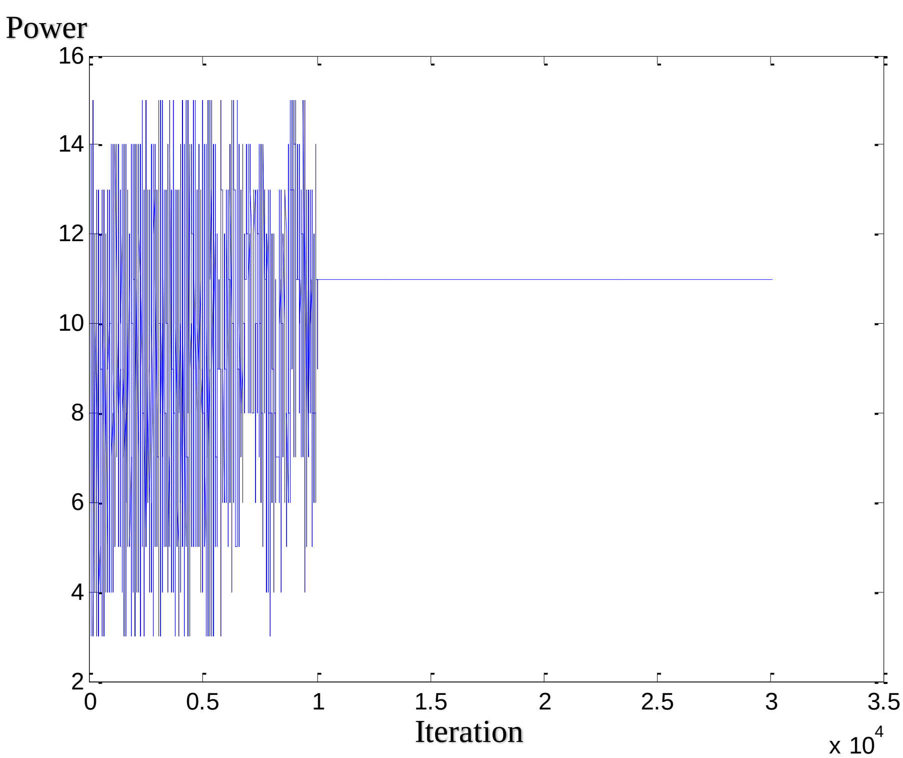
(1)


Power
16
14
12
10
8
6
4
2
0
0.5
1
1.5
2
2.5
3
3.5
x 10^4
Iteration
(2)

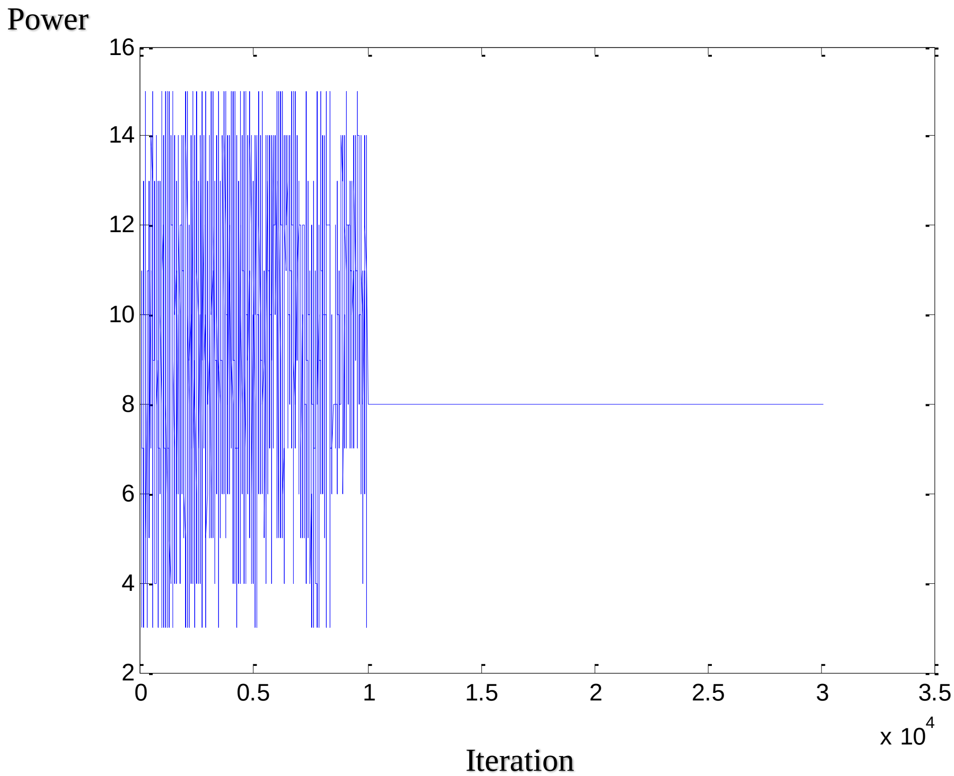

Power
20
18
16
14
12
10
8
6
4
0
0.5
1
1.5
2
2.5
3
3.5
x 10^4
Iteration
(3)


Power
16
14
12
10
8
6
4
2
0
0.5
1
1.5
2
2.5
3
3.5
Iteration
x 10^4

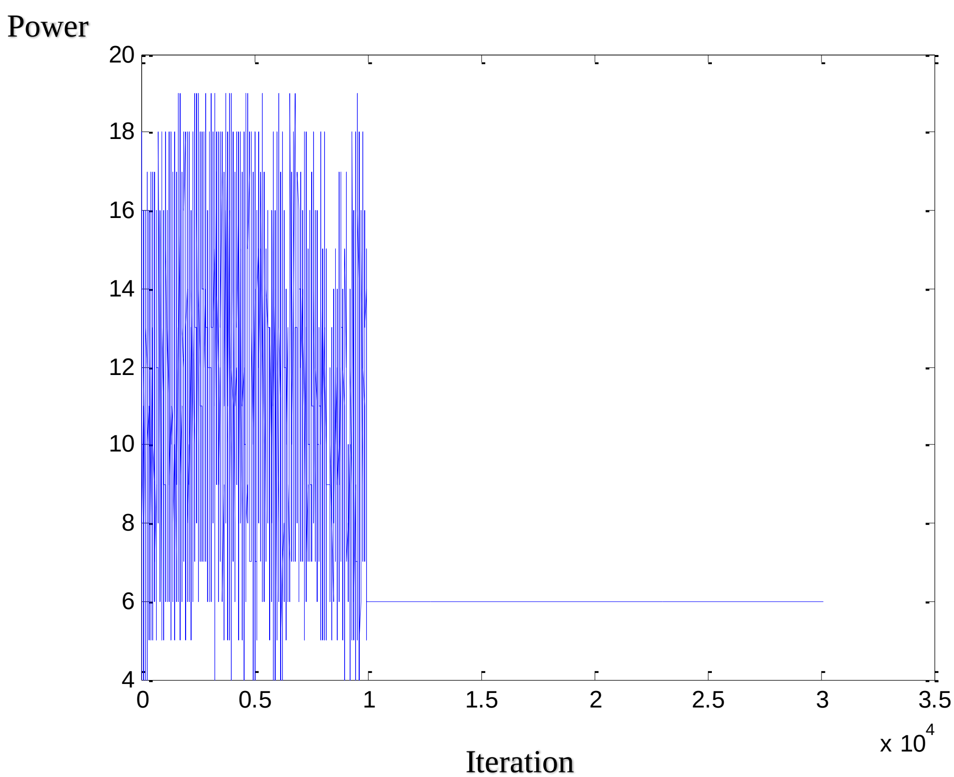
(4)


Power
16
14
12
10
8
6
4
2
0
0.5
1
1.5
2
2.5
3
3.5
Iteration
x 10^4

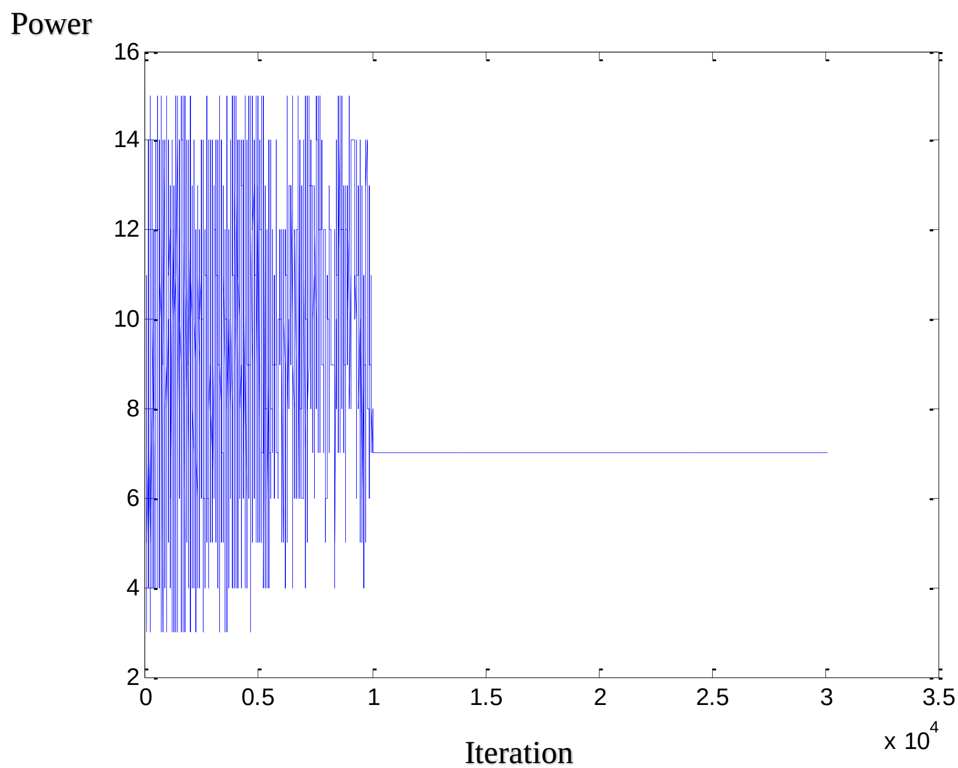
(5)



(6)



(7)

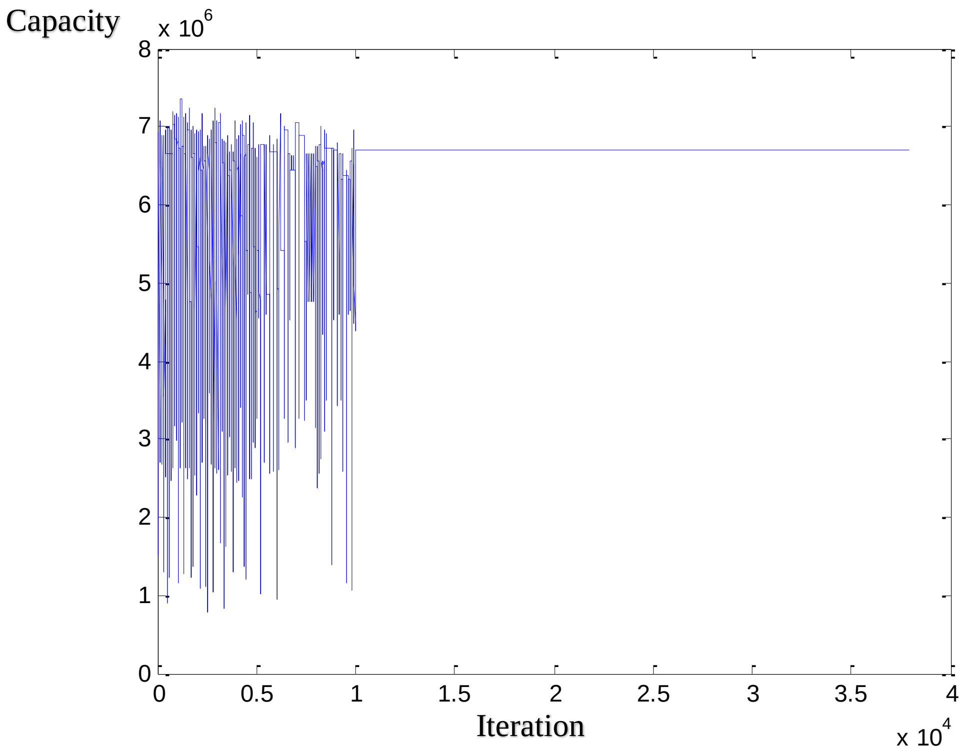


(8)

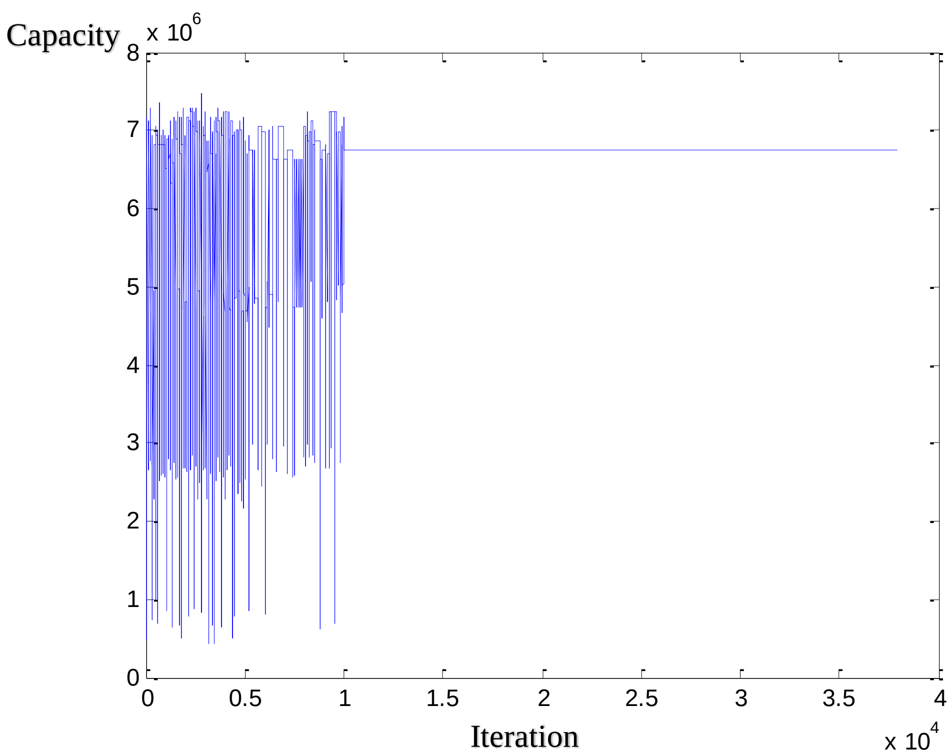
Figure 3.4: Femtocell power as Function of Learning Iterations

3.3.2 Scenario-2 Results

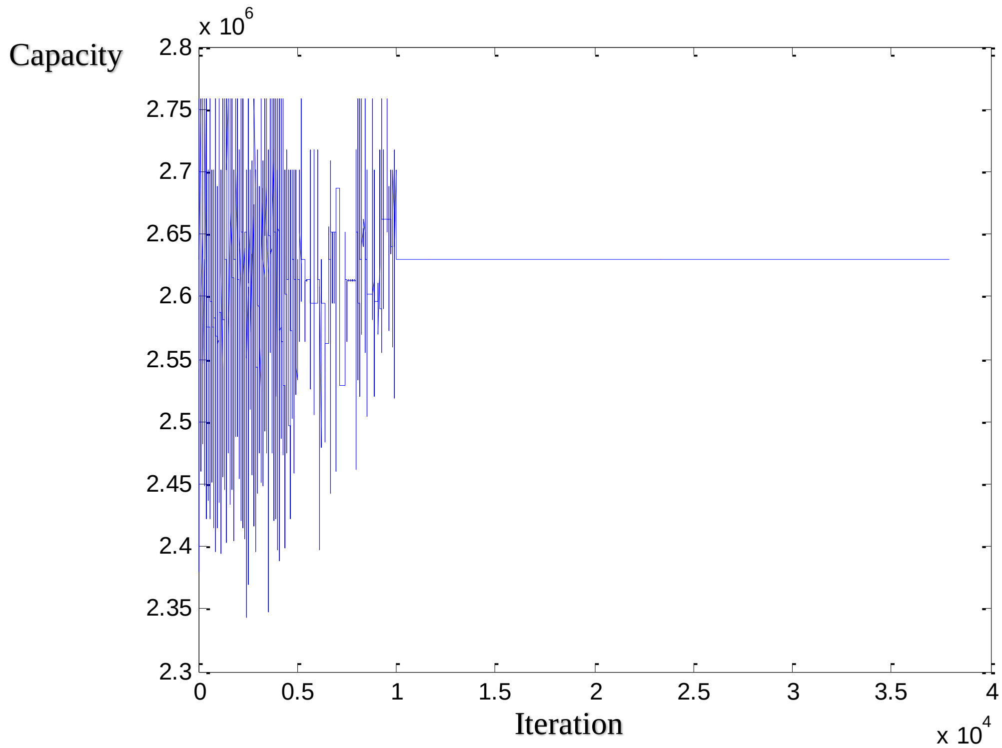
As in Scenario-1 results, results presented in this section correspond to a learning process modeled following the case study 2 with an ε-greedy constant action selection policy. As we mentioned, it has to be noted that the decentralized Q-learning algorithm, as any other learning scheme, needs a learning phase to learn the optimal decision policies. However, in this scenario Q-learning algorithm applied on femtocell to find the optimal resource allocation that ensure to serve efficiently the own users without attacking on other femto/macro users. Figure 3.5 show the capacity of overall system, 2- HeNBs and MeNB, as a function of the learning iterations. They can be observed that each HeNB finds optimal resource allocation for own users. So, the optimal actions ensure the efficient serves for users and do not exceed on the other users registered on different HeNB. Moreover, figure 3.6 show the power of the HeNBs as a function of learning iterations. Therefore, they can be observed, as we mentioned after each HeNB find optimal actions, the total power estimated to serve user without exceeding to maximum power threshold.



(1) HeNB-1

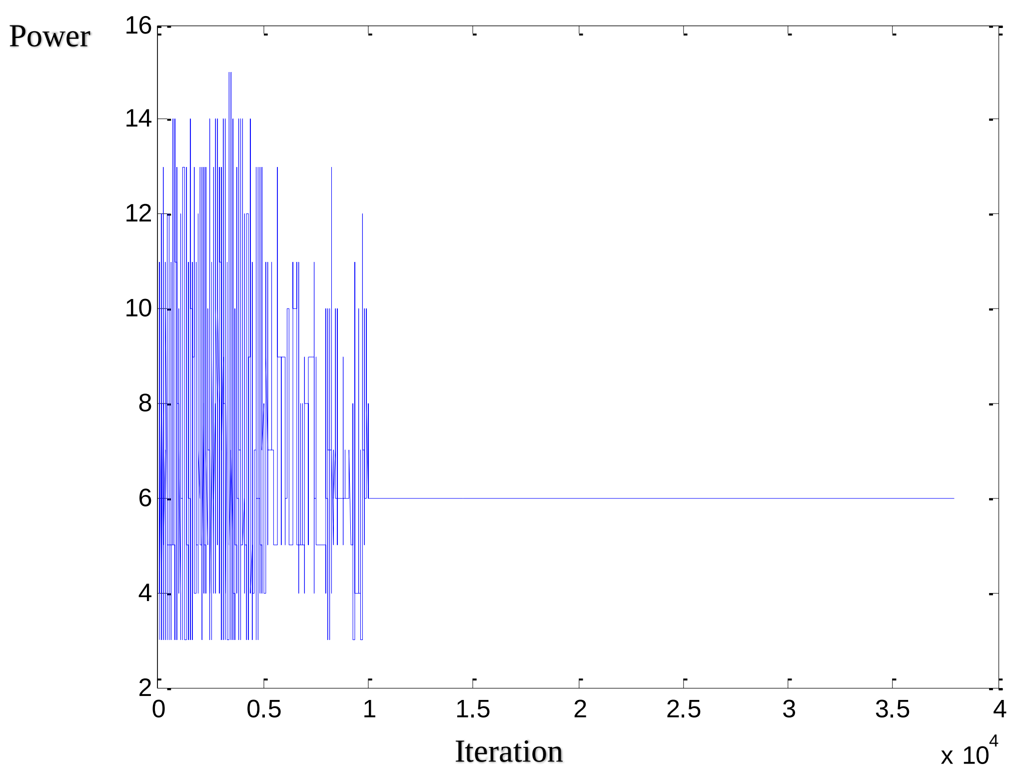


(2) HeNB-2

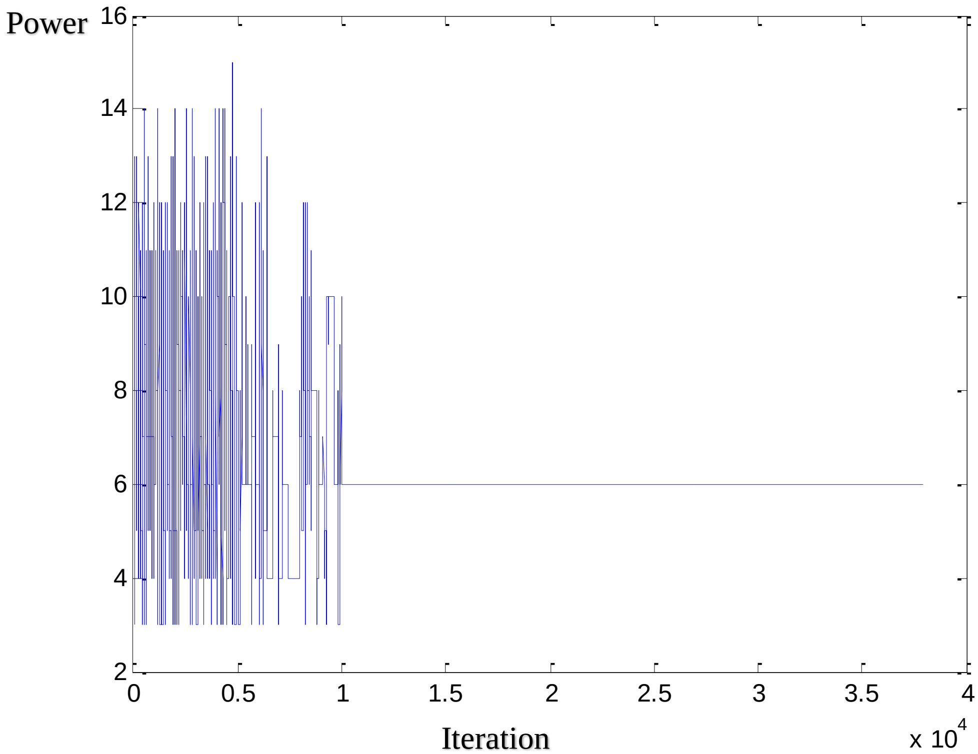


(3) MeNB

Figure 3.5: Femtocell/Macrocell Capacity as Function of Learning Iterations



(1) HeNB-1



(2) HeNB-2

Figure 3.6: Femtocell power as Function of Learning Iterations

The Simulation results have shown on figures (3.3-3.4) and (3.5-3.6) that each femtocell reach to optimality of resource block and power allocation in both scenarios, by applied machine learning (q-learning) capabilities.

Chapter Four

CONCLUSION AND FUTURE WORK

4.1 Conclusion

An overview of the OFDM technique was mentioned where we focused on the downlink part of system where the mean characteristics are viewed in details. Also, LTE architecture network was presented. Moreover, the physical link explains in details and showing interference problem in detail, and the importance of mitigation solutions in smooth service provisioning. Then, we sort out of femtocell network and its advantages for enhancement indoor coverage. After that, the obstacles facing the overall system were explained.

The propagation model is highlighted based on large and small scale fading where the mean fading are flat, frequency selective, fast and slow fading. Finally, a description of the machine learning characteristics was introduced; reinforcement learning (Q-learning) where the core elements of reinforcement were reminded in details. Moreover, the characteristic and summarize mathematics model of the agent leaning explained.

The main contribution of this thesis was presented and description the overall system then proposed the scenarios of the simulation. Thus, we modeled the decentralized Q-learning algorithm based on the theory of single agent learning to deal with the problem of interference generated by multiple femtocells which a joint resource allocation problem in femtocells system has been investigated where power allocation, resource block were jointly optimized. The main optimization problem is formulated in order to maximize the total end-to-end throughput of the system subject to interference constraints. The Q-learning approach is applied to solve the self-organization problem and achieve an optimal solution. Due to the autonomous femtocell configuration upon environment system , each femtocell reach to optimality that done to objectives those are to ensure giving efficient serves and mitigating the interference to other users.

4.2 Future Work

The work presented in this thesis left multiple investigation lines open for future work. In what follows we summarize the most important ones.

- The solutions presented in this thesis have been stated from a decentralized point of view, given the important advantages of this form of modeling. Nevertheless, it lacks a comparison with a centralized solution in order to have a quantitative measure of the gains introduced by the decentralization in the RRM procedures. Centralized solutions commonly have better performance given their complete knowledge about the system. However, they are difficult to implement, require large signaling and can incur in important delays. On the other hand, decentralized systems are fast and require little signaling, but can result in sub-optimal behaviors, given their partial knowledge about the system and usually show a slow convergence time.

- In this thesis it doesn't touching some tunable parameters in the learning algorithm i.e. learning rate and learning period, as well as a smart Q-table initialization, so it is recommended that check the gains of this parameters in terms of speed of convergence and accuracy.

- Another interesting point to be fulfilled is the combination of Fuzzy control with Q-learning which would allow to achieve a completely autonomous learning algorithm based on fuzzy control with better performances in terms low memory need where introduced by FIS instead of Q-table which more data from each agent respects to state-action need to saved.