LEARNING-BASED MECHANISMS TO ENHANCE LTE-U OPERATION

A Degree Thesis

Submitted to the Faculty of the

Escola Tècnica d'Enginyeria de Telecomunicació de Barcelona

Universitat Politècnica de Catalunya

by

Magí Coscollola Coletes

In partial fulfilment

of the requirements for the degree in

SCIENCE AND TELECOMMUNICATION TECHNOLOGIES ENGINEERING

Advisor: Oriol Sallent Roig

Barcelona, July 2015

Abstract

Nowadays, mobile communications are evolving to more complex systems and environments, LTE-U is a clear example as it needs to work in the unlicensed spectrum, which is a highly occupied and changing environment. A good solution for the channel selection in this case is using learning mechanisms. Learning techniques provide the systems with the ability to learn from the environment in order to make the best decisions.

In this thesis you can find the evolution of the LTE technology until arriving to LTE-U, and also a brief explanation about learning.

The project also explores evolutions of a given channel selection algorithm based in Q-learning, firstly finding initial values that make it work better and, secondly, proposing some variations in order to improve its performance.

Resum

Avui en dia, les comunicacions mòbils estan evolucionant cap a sistemes i entorns més complexos, l'LTE-U n'és un clar exemple ja que ha d'operar en l'espectre lliure, que és un entorn molt ocupat i constantment en canvi. Una bona solució per la selecció de canal és, en aquest cas, fer servir mecanismes d'aprenentatge. Les tècniques d'aprenentatge donen als sistemes l'habilitat d'aprendre de l'entorn per poder prendre les millors decisions possibles.

En aquesta tesi podreu trobar l'evolució de la tecnologia LTE fins arribar a l'LTE-U, i també una breu explicació sobre l'aprenentatge.

El projecte també explora evolucions d'un algorisme de selecció de canal basat en *Q-learning* lliurat prèviament, primer trobant valors inicials que el fan treballar millor i, després, proposant algunes variacions per augmentar el seu rendiment.

Resumen

Hoy en día, las comunicaciones móviles están evolucionando hacia sistemas y entornos más complejos, LTE-U es un claro ejemplo ya que tiene que operar en el espectro libre, que es un entorno altamente ocupado y en constante cambio. Una buena solución para la selección de canal en este caso es el uso de mecanismos de aprendizaje. Las técnicas de aprendizaje proveen a los sistemas la habilidad de aprender de su entorno para poder tomar las mejores decisiones posibles.

En esta tesis podréis encontrar la evolución de la tecnología LTE hasta llegar a LTE-U, y también una breve explicación sobre el aprendizaje.

El proyecto también explora evoluciones de un algoritmo de selección de canal basado en *Q-learning* librado previamente, primero encontrando valores iniciales que lo hacen trabajar mejor y, después, proponiendo algunas variaciones para aumentar su rendimiento.

Acknowledgements

I would like to express my gratitude to the project supervisor, Oriol Sallent, for his guidance and help during all the semester, and also for accepting to be my supervisor even though he has been living and working in Ireland during this time.

I also give thanks to Sergio Garcia for registering me in the servers of the D4 building.

I would also like to thank all the people in the GRCM for their publications, they have been very useful to learn about this area.

Finally, I thank my friends and family for their unconditional support.

Revision history and approval record

Revision	Date	Purpose
0	17/06/2015	Document creation
1	22/06/2015	Document revision
2	30/06/2015	Document revision
3	03/07/2015	Document revision
4	06/07/2015	Document revision
5	08/07/2015	Document revision

DOCUMENT DISTRIBUTION LIST

Name	e-mail
Magí Coscollola Coletes	mcoscollola@gmail.com
Oriol Sallent Roig	sallent@tsc.upc.edu

Written by:		Reviewed and approved by:	
Date	10/07/2015	Date	10/07/2015
Name	Magí Coscollola Coletes	Name	Oriol Sallent Roig
Position	Project Author	Position	Project Supervisor

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

telecom BCN

Table of contents

Abstract 1
Resum 2
Resumen 3
Acknowledgements 4
Revision history and approval record 5
Table of contents 6
List of figures 8
List of tables: 9
List of mathematical expressions: 10
1. Introduction 11
  1.1. Project background 11
  1.2. Project requirements and specifications 11
  1.3. Work plan 12
    1.3.1. Work Packages, Tasks and Milestones 12
      1.3.1.1. Work packages: 12
      1.3.1.2. Milestones 14
    1.3.2. Gantt diagram 15
    1.3.3. Deviations from the initial plan 15
2. State of the art of the technology used or applied in this thesis: 16
  2.1. LTE technology [2][3][4] 16
    2.1.1. LTE Release 8 Requirements 16
    2.1.2. Release 9 [5][6] 17
  2.2. LTE-A[7][8][9] 18
    2.2.1. LTE-A requirements 18
    2.2.2. LTE-A Release 10 features 19
    2.2.3. Release 11 [10][11] 19
    2.2.4. Release 12 [12][13] 21
  2.3. LTE-U[14] 22
    2.3.1. Band definitions for LTE-U 23
    2.3.2. Channel selection in LTE-U 23
    2.3.3. Channel Access in LTE-U [1] 24
  2.4. Artificial intelligence and learning algorithms[17][18] 24
    2.4.1. Machine learning[18][19] 25
    2.4.2. Q-learning 25
      2.4.2.1. Q-learning-based channel selection in LTE-U[1] 26

2.5. Simulation scenario and code 27
2.5.1. Simulation scenario 27
2.5.2. Simulator code 28
3. Project development and results: 31
3.1. Study of the impact of the initial value of Q on the convergence time of the system 31
3.1.1. Environment 31
3.1.2. Simulation results 31
3.2. Study of the impact of the number of channels on the convergence time of the system 36
3.2.1. Environment 36
3.2.2. Simulation results 37
3.3. A new mechanism to improve the convergence time of the system: Semi-fixed Q-learning 38
3.3.1. Coding 39
3.3.2. Environment 40
3.3.3. Simulation results 40
3.4. Semi-fixed Q-learning version 2 41
3.4.1. Environment 41
3.4.2. Simulation results 41
4. Budget 49
5. Conclusions and future development: 50
Bibliography: 51
Glossary 53

List of figures

Figure 1. Unlicensed spectrum availability in different regions.[16] 23

Figure 2. Illustration of the two coexistence mechanisms 24

Figure 3. Distribution of the small cells in the simulator scenario. 27

Figure 4. Example of the channel selection probability plots. 32

Figure 5. Evolution of the convergence time in function of Q_0. 34

Figure 6. Comparison between the fourth group of simulations and the mean of the other three groups. 34

Figure 7. Q evolution when $Q_0 = 0.1$. 35

Figure 8. Q evolution when $Q_0 = 0.7$. 36

Figure 9. Q evolution of AP3 in simulation 700 with 4 available channels. 38

Figure 10. Distribution of the APs. 39

Figure 11. Evolution of Q in SC1 using Semi-fixed Q-learning v2 42

Figure 12. Evolution of channel selection probability in SC1 using Semi-fixed Q-learning v2. 42

Figure 13. Evolution of the throughput in SC1 using Semi-fixed Q-learning v2. 42

Figure 14. Evolution of current channel in SC1 using Semi-fixed Q-learning v2. 42

Figure 15. Q evolution when one interference has been included. 44

Figure 16. Channel selection probability evolution when one interference has been included. 45

Figure 17. Q evolution with both default channels of SC1 interfered. 46

Figure 18. Q evolution with both default SP1 channels interfered and reward type ‘total’. 47

List of tables:

Table 1: Data rate and spectrum requirements defined for LTE 17
Table 2. Results of the first group of simulations..32
Table 3. Results of the second group of simulations..33
Table 4. Results of the third group of simulations. 33
Table 5. Results of the convergence time calculation with 4 or 8 available channels. 37
Table 6. Convergence time using Semi-fixed Q-learning v2...43
Table 7. Total software cost...49
Table 8. Total hour cost...49

List of mathematical expressions:

(1) Q-learning update expression........26
(2) Reward formula........26
(3) Throughput served by a cell i using the channel k........26
(4) Channel selection probability........27
(5) Cooling function........27
(6) Channel selection probability formula........39
(7) Approach of the denominator of (6) to a constant........39
(8) Deduction using (6) and (7)........40
(9) Formula to calculate r_{TOTAL}........47
(10) Formula to calculate $r_{OPERATOR}$........47

1. Introduction

This project has been carried out at the Mobile Communications Research Group (GRCM), in the Department of Signal Theory and Communications (TSC) from the UPC.

This project consists on analyzing and working with a Matlab code which simulates a LTE-U environment in order to improve the downlink channel selection algorithm of the access points. In this way we provide a good tool for LTE-U base stations to avoid interferences between them and also with other services that are using the 5 GHz unlicensed band.

The documentation part is also important in this project. It existed the need to learn about LTE technology and learning algorithms, and why they have 'met' each other.

The main goal of this project is develop and program a better channel selection algorithm than the one from our start point and prove it using simulation results.

1.1. Project background

LTE-U (Long Term Evolution Unlicensed) is an enhancement in LTE that enables it to operate and coexist with other technologies in unlicensed bands. The use of the unlicensed spectrum in addition to the LTE licensed spectrum will be an important complement to meet the ultra-high capacity that will be needed for 4G and beyond.

The introduction of LTE-U brings a number of challenges to be addressed. As unlicensed spectrum, LTE-U must support fair access of multiple LTE-U networks and multiple Wi-Fi networks. When demand exceeds capacity, each network should be able to access an equal share. When a particular network's traffic demand is less than the spectral capacity of an equal share, that network should allow other networks to access the unused capacity. This will require LTE-U to adapt to the presence of other LTE-U and Wi-Fi networks, while Wi-Fi uses its current mechanisms. Therefore, issues such as coexistence with Wi-Fi systems operating in unlicensed spectrum, unpredictable interference to LTE from other technologies and coexistence among cells from the same or different operators need to be resolved.

This project has its initial point in a draft of a paper that was going to be published by the GRCM [1] (in the moment of starting the project the paper has not been published) and the Matlab code associated to it. The paper focused on the channel selection algorithm in LTE-U environments in order to decide the most appropriate channel in the unlicensed band to set-up a LTE-U carrier. The algorithm they proposed was based in Q-Learning and independent access points.

1.2. Project requirements and specifications

Project requirements:

- LTE-U has to be a 'good neighbor'.
- The algorithm has to work trying to satisfy the LTE-U demand.
- All the access points have to get a channel to operate.
- It has to work in indoor and outdoor placements.

Project specifications:

- It must satisfy the 3GPP specifications.
- The algorithm has to do a Channel Availability Check in order to detect the presence of radar signals. Only channels with no radar signals can be selected for operation.

- To ensure a fair coexistence in the time domain in the same channel, a Listen-Before-Talk scheme operating at milliseconds scale is needed.
- The transmission should be switched to a cleaner channel using LTE Rel. 10/11 procedures if there are interferences in the current one.

1.3. Work plan

1.3.1. Work Packages, Tasks and Milestones

1.3.1.1. Work packages:

Project: Learning-based mechanisms to enhance LTE-U operation	WP ref: 1	
Major constituent: Read documentation	Sheet 1 of 5	
Short description: Read the necessary documentation in order to have enough knowledge to start with the project.	Planned start date: 17/2/15 Planned end date: 27/2/15 Start event: End event:	
Internal task T1: Read LTE-U paper Internal task T2: Read bibliography	Deliverables:	Dates:

Project: Learning-based mechanisms to enhance LTE-U operation	WP ref: 2	
Major constituent: Matlab simulator	Sheet 2 of 5	
Short description: Practice with the simulator in order to get used to it and be able to make future modifications.	Planned start date: 28/2/15 Planned end date: 8/3/15 Start event: End event:	
Internal task T1: Get used to Matlab code	Deliverables:	Dates:

Project: Learning-based mechanisms to enhance LTE-U operation	WP ref: 3	
Major constituent: Create new algorithms	Sheet 3 of 5	
Short description: Find points that can be modified in the algorithm and write new solutions in order to improve the throughput of the system. After the two first tasks starts WP4 and WP3 is used to improve the solutions in WP4.	Planned start date: 9/3/15 Planned end date: 14/6/15 Start event: End event:	
Internal task T1: Find points to improve Internal task T2: Create the improvements Internal task T3: Improve the algorithms with the results in WP4	Deliverables:	Dates:

Project: Learning-based mechanisms to enhance LTE-U operation	WP ref: 4	
Major constituent: Program and test	Sheet 4 of 5	
Short description: Program the improvements written in WP5 and test their performance.	Planned start date: 25/4/15 Planned end date: 14/6/15 Start event: End event:	
Internal task T1: Program the new algorithm Internal task T2: Test and get the results	Deliverables:	Dates:

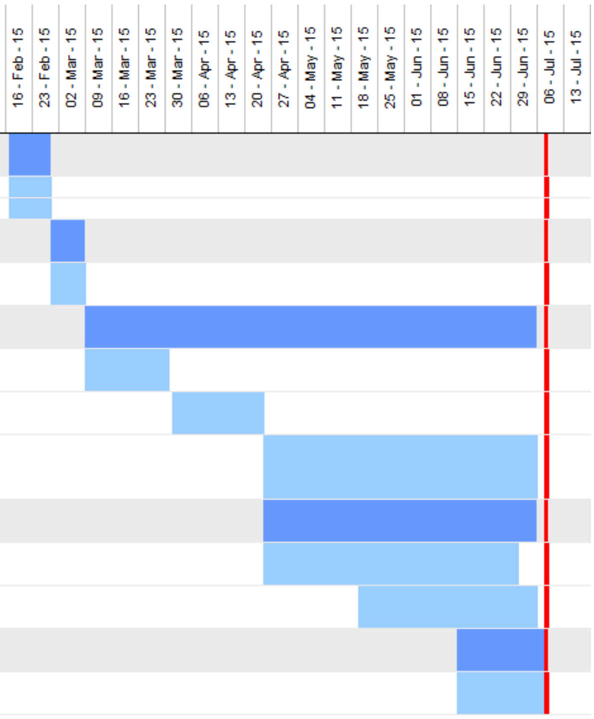
Project: Learning-based mechanisms to enhance LTE-U operation	WP ref: 5	
Major constituent: Conclusions and report	Sheet 5 of 5	
Short description: Collect all the results and write the final report of the project.	Planned start date: 15/6/15 Planned end date: 30/6/15 Start event: End event:	
Internal task T1: Collect all the results write the report	Deliverables:	Dates:

1.3.1.2. Milestones

WP#	Task#	Short title	Milestone / deliverable	Date (week)
1	1.1	Read LTE-U paper		27/2/15
	1.2	Read bibliography		27/2/15
2	2.1	Get used to the Matlab code	**Project Proposal and Work Plan**	6/2/15
			Have got used to the code	8/2/15
3	3.1	Find points to improve	Have found some points to improve	30/3/15
	3.2	Create the improvements	Have written in pseudocode the improvements	24/4/15
3 & 4	3.3		**Critical Review**	24/4/15
4	4.1	Program the new algorithm	Have the new algorithm working properly	19/5/15
	4.2	Test and get the results		14/6/15
5	1	Collect all results and compare	Have collected all the results and made the comparison	30/6/15

1.3.2. Gantt diagram

WBS	Tasks	Task Lead	Start	End	Duration (Days)	% Complete	Working Days	Days Complete	Days Remaining
1	**READ DOCUMENTATION**	Magi	2/17/15	2/27/15	11	100%	9	11	0
1.1	Read LTE-U paper		2/17/15	2/27/15	11	100%	9	11	0
1.2	Read bibliography		2/17/15	2/27/15	11	100%	9	11	0
2	**MATLAB SIMULATOR**	Magi	2/28/15	3/8/15	9	100%	5	9	0
2.1	Get used to the Matlab code		2/28/15	3/8/15	9	100%	5	9	0
3	**CREATE NEW ALGORITHMS**	Magi	3/9/15	7/5/15	119	100%	85	119	0
3.1	Find points to improve		3/9/15	3/30/15	22	100%	16	22	0
3.2	Create the improvements		4/1/15	4/24/15	24	100%	18	24	0
3.3	Improve the algorithms with the results in WP4		4/25/15	7/5/15	72	100%	50	72	0
4	**PROGRAM AND TEST**	Magi	4/25/15	7/5/15	72	100%	50	72	0
4.1	Program the new algorithms		4/25/15	6/30/15	67	100%	47	67	0
4.2	Test and give feedback to WP3		5/20/15	7/5/15	47	100%	33	47	0
5	**CONCLUSIONS AND REPORT**	Magi	6/15/15	7/8/15	24	100%	18	24	0
5.1	Collect all the results and write the report		6/15/15	7/8/15	24	100%	18	24	0



1.3.3. Deviations from the initial plan

Due to the need of more time to do the simulations to find the optimal Q_0, as we have written before, we had to change several dates from the initial plan in the moment of the critical design review:

- WP3 – 3.2 Create the improvements: End date modified from 20/4/2015 to 24/4/2015.
- WP4 – 4.1 Program the new algorithm: Start date modified from 21/4/2015 to 25/4/2015. End date modified from 10/5/2015 to 19/5/2015.
- WP4 – 4.2 Test and give feedback to WP3: Start date modified from 11/5/2015 to 20/5/2015.

After this critical design review, we changed the following things:

- WP3 – 3.3 Improve the algorithms with the results in WP4. End date modified from 14/6/2015 to 5/7/2015, because we have not finished in the moment of starting to write the report, so we had perform the work packages in parallel.
- WP4 – 4.1 Program the new algorithms. We decided to program another algorithm to improve the 'new' one, so we had to change the end date from 19/5/2015 to 30/06/2015.
- WP4 – 4.2 Test and give feedback to WP3. Due to the previous change, we also had to change this point. We moved the end date from 14/6/2015 to 5/7/2015.
- WP5 – 5.1 Collect all the results and write the report. As we had to work in parallel with WP4, we had to move the end date from 30/6/15 to 8/7/15.

2. State of the art of the technology used or applied in this thesis:

2.1. LTE technology [2][3][4]

Long Term Evolution (LTE) is the technology that was chosen in 2008 in order to become the fourth generation of mobile communications. It presented a big difference with the previous technologies because LTE is an “all-IP”-based network, so the information is always sent using packet switching while the previous generations had both circuit and packet switching modes.

The studies in this technology started in November 2004, when the term E-UTRAN[1] was created as an evolution of UTRAN. One month later, the study item “Evolved UTRA and UTRAN” was created by 3GPP in order to evolve to a new technology with higher transmission rates and lower latency and optimized for packet scheduling.

Some of the objectives of E-UTRA were:

- Reduced cost per bit.
- Increased service provisioning – more services at lower cost with better user experience.
- Flexibility of use of existing and new frequency bands.
- Simplified architecture, open interfaces.
- Allow for reasonable terminal power consumption.
- Transmission rates of 100 Mbps in downlink and 50 Mbps in uplink.
- Increase the spectral efficiency in a factor between 2 and 4 from Release 6.
- Time between the IP layer in the device and the IP layer in the radio access network (or vice versa) lower than 30ms.
- Scalable bandwidth.
- Interoperability with 3G and non 3GPP systems.

Finally, in December 2007, the first version of the LTE specifications was approved and in 2008 Release 8 was published.

In order to reach these objectives, LTE was based on new technical principles. It used a new multiple access scheme: OFDMA in downlink and SC-FDMA in uplink. They used this multiple access mechanism in the uplink because OFDMA properties are less favourable for it, while they are optimum for the downlink. Another essential part of LTE are MIMO antenna schemes. LTE also includes an FDD operation mode and a TDD operation mode.

2.1.1. LTE Release 8 Requirements

LTE fulfilled most of the initial requirements and exceeded some of them:

- **Data Rate:** Peak rates target 100 Mbps in the downlink and 50 Mbps in the uplink, for 20 MHz bandwidth and assuming 2 receiver and 1 transmitter antennas.
- **Throughput:** Downlink average user throughput per MHz is 3-4 times better than 3GPP Rel. 6. Uplink average user throughput per MHz is 2-3 times better than 3GPP Rel. 6.
- **Spectrum efficiency:** Downlink target is 3-4 times better than 3GPP Rel. 6. Uplink target is 2-3 times better than 3GPP Rel. 6.

[1] The terms LTE, E-UTRA or E-UTRAN are used interchangeably.

Downlink (20 MHz)			Uplink (20 MHz)		
Unit	**Mbps**	**Bps/Hz**	**Unit**	**Mbps**	**Bps/Hz**
Requirement	100	5	**Requirement**	50	2.5
2x2 MIMO	172.8	8.6	**2x2 MIMO**	57.6	2.9
4x4 MIMO	326.4	16.3	**4x4 MIMO**	86.4	4.3

Table 1: Data rate and spectrum requirements defined for LTE

- **User plane latency:** less than 30 ms between a packet being at the IP layer in the device and the packet being at the IP layer in the radio access network, and vice versa.
- **Control plane latency:** less than 100 ms between the device is in 'idle' state and the device is in 'connected' state.
- **Control plane capacity:** at least 200 users per cell should be supported in the active state for spectrum allocations up to 5 MHz.
- **Bandwidth:** LTE supports a subset of bandwidths of 1.4, 3, 5, 10, 15 and 20 MHz.
- **Mobility:** the system shall be optimized for low mobile speed (0-15 km/h). Also higher speeds shall be supported, including high speed train environment.
- **Spectrum allocation:** operation in paired (FDD) and unpaired spectrum (TDD) is possible.
- **Co-existence:** co-existence with 2G and 3G shall be ensured, and also co-existence between multiple operators.
- **Coverage:** throughput, spectrum efficiency and mobility targets above should be met for 5 km cells, and with a slight degradation for 30 km cells.
- **Multimedia Broadcast Multicast Services (MBMS):** provision of simultaneous dedicated voice and MBMS services to the user, available for paired and unpaired spectrum arrangements.
- **Radio Resource Management:** enhanced support for end to end QoS, efficient support for transmission of higher layers and support of load sharing and policy management across different Radio Access Technologies (RAT).

2.1.2. Release 9 [5][6]

3GPP Release 9 was published in 2009. The new release included a list of features that were not completed in Release 8. These are:

- **MBMS:** they provide to the operators the possibility to broadcast multimedia content over the existing cellular network. In LTE it offers 20 mobile TV channels at 256 kbps in a 5 MHz channel.
- **LTE MIMO:** one of the technologies that helps to get high data rate, high system capacity and large coverage is Beamforming. It improves the cell edge performance. In Rel-8 LTE supports single-layer beamforming, it is based on user-specific Reference Symbols. In general, this solution allows to direct the beam towards a specific UE through position estimation at the eNB. It is especially suited for LTE in TDD mode. To further evolve this technology, in Release 9 we find dual-layer beamforming.
- **LTE positioning:** Today most of modern mobile devices have an integrated GNSS (Global Navigation Satellite Systems) receiver which allows them to know their exact position. The problem is that the receiver needs to have an unobstructed line of sight to at least four satellites, and this is impossible indoors or on some urban environments. To overcome this situation, Assisted-GNSS has been developed.

This technology uses network resources to provide assistance data that helps the device to locate itself. In LTE additional methods have been standardized and existing ones, enhanced.

- **Public Warning System:** the LTE network can be used to broadcast warning notifications in case of emergencies.
- **RF requirements for multi-carrier and multi-RAT base stations:** multi RAT base stations are those that have LTE and other technologies like UMTS or GSM.
- **Home eNodeB[2] specification (femto-cell):** Release 9 introduces some new requirements in order to provide the LTE experience for user through femto-cells.
- **Self-Organizing Networks:** it is a system that allows the base stations change their settings in order to optimise their operating expenses and the network quality.
- **LTE Pico NodeB requirements:** it introduces the RF requirements for LTE Pico BS. Pico BS has small coverage and is typically used to increase capacity in areas with dense user population and high traffic intensity.

2.2. LTE-A[7][8][9]

Even though LTE supposed a notable improvement in that moment's technology, it was not considered 4G by the ITU, because it did not achieve the 4G requirements. LTE is called 3.9G because it is close to 4G requirements but it doesn't reach them.

However, the evolution of LTE, LTE Advanced, is a technology which became a recognised 4G technology by the ITU. This evolution came with 3GPP Release 10 in 2009, and got the acknowledgement in October 2010.

2.2.1. LTE-A requirements

The requirements for LTE-A were based on the requirements for IMT-Advanced systems, issued by the ITU for what is commercialised as 4G.

LTE-A requirements are listed below, and in general they have been reached or even exceeded.

- **Peak data rate:** the system should target a 1 Gbps downlink data rate and a 500 Mbps uplink data rate.
- **Latency:** in control plane it should be less than 50 ms to change from idle mode to connected mode. In user plane the latency should be less than in LTE Release 8.
- **Spectrum efficiency:** it aims to support downlink (8x8 antenna configuration) peak spectrum efficiency of 30 bps/Hz and uplink (4x4 antenna configuration) peak spectrum efficiency of 15 bps/Hz.
- **Cell edge user throughput:** LTE-A should allow it to be as high as possible. It is defined as the 5% point of the CDF of the user throughput normalized with the overall cell bandwidth.
- **VoIP capacity:** it should have been improved in comparison to Release 8.
- **Mobility:** the system shall support mobility for various mobile speeds up to 350 km/h. The system performance shall be enhanced for 0 up to 10 km/h.
- **Spectrum flexibility:** LTE-A shall operate in spectrum allocations of different sizes including wider spectrum allocations than the ones of Release 8. FDD and TDD should be supported for existing paired and unpaired frequency bands, respectively.

[2] eNodeB: Evolved NodeB. Is how base stations are called in LTE. It takes the name 'NodeB' from UMTS.

2.2.2. LTE-A Release 10 features

LTE-A includes some new features that weren't in LTE or extends some of the LTE features. Several of this characteristics are listed below:

- **Band aggregation:** a solution to reach high data rates requirements is to use more bandwidth, i.e. aggregate multiple LTE carriers. Two or more component carriers are aggregated in order to support bandwidths up to 100 MHz.
- **Enhanced multiple antenna technologies:** LTE-A extends the MIMO capabilities of Release 8 to four uplink and eight downlink layers.
- **Enhanced uplink transmission scheme:** it has been maintained the SC-FDMA system, but some improvements have been done, like decoupling of control information and data transmission.
- **Multi-cluster transmission:** With SC-FDMA the transmission in the uplink is always contiguous, the terminal transmits only on consecutive subcarriers. With LTE-A it has been introduced clustered SC-FDMA, i.e., the uplink transmission is not anymore restricted to the use of consecutive subcarriers.
- **Enhanced Inter-cell Interference Coordination (eICIC):** as LTE is a single frequency network, the management of interferences on cell borders was an important topic. With the appearance of Heterogeneous Networks[3] the methods used in Release 8 and 9 are not sufficient anymore, so some new mechanisms were needed to be created.
- **Relaying:** LTE-A extends LTE with support for relaying in order to enhance coverage and capacity, creating Heterogeneous Networks, so while an UE is connected directly to the eNodeB, another can be communicating with a relay node which communicates with a donor eNodeB.
- **LTE Self Optimizing Networks enhancements:** it continues the work started in Rel-9. Capacity and Coverage Optimization, to enable the detection of coverage and capacity problems; Mobility Robustness Optimization enhancements, to enable the detection and provide tools for possible correction of connection failures, unsuccessful re-establishments after connection failure, ping-pong handovers[4] and handovers to wrong cells; Mobility Load Balancing enhancements and energy savings are introduced.

2.2.3. Release 11 [10][11]

3GPP Release 11 started in 2009 and got frozen in June 2013. Some of the enhancements we can find in this release are in the next list:

- **Carrier Aggregation:** it was the most demanded feature in Release 10 due to its capability to sum up the likely fragmented spectrum that a network operator owns. In Rel-11 we find some enhancements of this technology. We can find multiple timing advances for uplink, in Rel-10 it had the same TA for all the component carriers so in Rel-11 it allows to work properly in scenarios where different delays are applied to each carrier. We also find non-contiguous intra-band carrier aggregation, which was not fully completed in Rel-10 even though it was mentioned in it; support of different UL/DL configurations in TDD mode, when in the previous release all carrier frequencies had to use the same UL/DL ratio, and the possibility to apply diversity in UL direction using the two antennas that the end user devices generally have, but it was decided that UL diversity can only be used if the UE is carried aggregation capable or configured with more than one cell.

---

[3] Heterogeneous Networks: networks built by a macro cell to ensure coverage and pico cells, femto cells and relay stations to illuminate shaded regions or to enhance data rate in hot spots.

[4] Ping-pong handover: as the name suggest, ping-pong phenomenon is when a UE performs two consecutive handovers between two different eNB, as if the UE was the ball in a ping pong match.

- **Coordinated Multi-point Operation:** CoMP is one of the most important improvements in Release 11. It shall allow the optimization of transmission and reception from multiple distribution points (multiple cells, for example) in a coordinated way. It reduces power consumption of the devices and improves overall throughput. It also reduces inter-cell interference.
- **E-PDCCH:** is a new downlink control channel necessary to support new features like CoMP or DL MIMO.
- **Further enhanced non CA-based ICIC:** feICIC enhances the Rel-10 eICIC providing the UE with Cell-specific Reference Signal (CRS) assistance information of the interfering cells in order to aid the UE to mitigate their interference. Rel-10 did not address CRS interference control and CRS must be still transmitted in order to ensure backward compatibility. CRS are used for cell search, initial acquisition and DL channel estimation and quality measurements.
- **Network Based Positioning:** improvements for the LTE positioning, adding the UL positioning. This method makes use of the measured timing at multiple base stations of UE signals to determine the UE's exact location. This improvement needs new equipment to be installed in the base stations, but it does not have any impact on the UE implementation.
- **Service continuity improvements for MBMS:** some improvements are introduced to enhance the service continuity. In the previous releases the UE has to search again for the current service if it does a handover, and the user perceives this as a service interruption. In Rel-11, some signalling has been added in order to allow the UE to immediately switch to the proper frequency and channel and avoid these search times.
- **Signalling/procedures for interference avoidance for In-Device Coexistence:** nowadays UEs contain several wireless technologies transmitting or receiving RF signals simultaneously. This situation causes In-Device Coexistence (IDC) interference. The solution specified in Rel-11 allows the UE to send an IDC indication to the base station if it cannot resolve the interference by itself. This should allow the base station to take appropriate measures, like changing the LTE carrier, among others.
- **Enhancements for Diverse Data Applications (EDDA):** they take in account the problem that different applications in the UE cause a small but frequent data traffic between the user and the network. The goal was to optimise user experience by allowing the UE to ask for a more power efficient mode of operation.
- **Minimization of Drive Test:** its goal is to get information of the current network from measurements taken by the UE. Combining these measurements with information from the RAN, network optimization can be done in an efficient way. Drive tests[5] shall be decreased and only necessary for measurements that are not available for a UE.
- **Network Energy Saving:** it's necessary to investigate possible network energy saving mechanisms to reduce CO_2 emission and operating expenses of mobile network operators. In Rel-11 a method that partly switched off eNBs, which cover the same area, when the capacity is not needed was enhanced to cover the inter RAT case. LTE cells providing additional capacity can be switched off when its capacity is not needed. The basic coverage in this case may be provided by other LTE, UMTS or GSM cells.

[5] Drive tests are tests made by the network in order to get information from it. Basically, they can be UE-based, when multiple end-user devices are hooked up to get a look at how they perform on its network, or benchmark testing, which includes UE that use other operators' networks. [22]

- **Relays for LTE:** support for relays has been specified in Rel-10. In Rel-11 the remaining issues (some transmitter and receiver requirements for access and backhaul[6]) are completed.

2.2.4. Release 12 [12][13]

This release is, nowadays, the last one that has been frozen. It started in 2010 and it finished in March of 2015.

Some of the enhancements that Rel-12 provides for LTE are the following:

- **Downlink MIMO:** it introduces enhancements in the Channel State Information, which enable the eNodeB to complete the delivery of data packets earlier thus improving the spectral efficiency.
- **Small Cells:** Rel-12 introduces several enhancements in their physical layer, like improving spectrum efficiency by increasing the transmission efficiency and reducing overhead, creating mechanisms to mitigate the interference (powering On/Off the small cells) or increasing the highest supported modulation from 64 QAM to 256 QAM. This release also focused on mobility robustness, reducing the handover signalling load, and also focused on improved per-user throughput and system capacity using dual connectivity (when a UE is capable of using radio resources from at least two different access points).
- **Proximity Services (ProSe):** In ProSe communications, UEs that are near each other communicate directly rather than through the cellular network. Release 12 focuses on enabling direct broadcast communication between public safety personnel when a network is unavailable (e.g. after a disaster).
- **UE Receiver enhancements:** this release includes a new category of UE receivers (NAICS) that increase the interference cancellation and suppression exchanging static cell configuration information between the neighbouring eNBs.
- **HetNet Mobility:** Release 12 provides means to improve overall handover performance in Heterogeneous Network environments. Optimal configuration of parameters and better speed estimation are seen as potential solutions. Faster re-establishments after a HO failure are introduced to reduce interruption time for the users and improve their experience.
- **Further enhancements for HeNB mobility:** it introduces an X2-Gateway for enhanced mobility procedures for LTE. In this way, a HeNB can connect to a peer (H)eNB using either direct X2[7] or through the X2-GW.
- **8 Rx Antennas for UL:** as the amount of uploaded data has increased due the use of intelligent terminals and new applications like social networking an improvement in the uplink was needed. Deployment of 8 reception antennas at the eNodeB is an efficient way to improve LTE UL performance in terms of capacity improvement, coverage extension and UE Tx power reduction.
- **MBMS:** there have been implemented mechanisms to re-establish the MBMS session after a failure, also measurements targeting MBMS Single Frequency Network signals are introduced in order to provide better tools for the network to monitor and adjust MBMS operational parameters. It also establishes that MBMS and Public Safety Services share the resources, so, when an emergency occurs, in its area Public Safety Services take MBMS resources.
- **Inter-eNB CoMP:** CoMP in Rel-11 did not address the specified support of a network interface for CoMP involving multiple eNBs with non-ideal backhaul. In Rel-12 this study has identified the cases that CoMP can provide a performance

[6] Backhaul: is the return network. In the case of relays is the link that connects the relay with the main base station.

[7] X2 is an interface that link eNBs with each other. Its main aim is to reduce the packet loss due to user mobility. This infrastructure has been implemented in LTE.

enhancement. Therefore enhancement on network interface and signalling messages should be specified. According to it, it should specify the signalling support based on X2 interface.

- **Enhanced Interference Management and Traffic Adaptation:** to better utilise spectrum in a TDD system, a TDD configuration that matches the traffic could be selected. This is the scope of this enhancement. To enable traffic adaptation, UEs are configured with a TDD configuration for the UL and a second one for the DL, whose subframes are dynamically selected by the network. The base station provides an indication to the UEs of what subframes will be used in each case.
- **FDD-TDD Carrier Aggregation:** within Release 12, 3GPP has specified support for allowing UEs to operate TDD and FDD spectrum jointly. The main solution to be specified is CA between the FDD and TDD spectrum. It would allow user throughputs to be boosted and it would allow a better way to divide the load in the network between FDD and TDD spectrum. Also dual connectivity between TDD and FDD is specified, which provides a tool to connect UEs to cells that are operating either TDD or FDD.

2.3. LTE-U[14]

LTE-U (Long Term Evolution Unlicensed) is an enhancement in LTE-A which enables it to operate and coexist with other technologies in unlicensed bands, it will be included in 3GPP Release 13 [15]. The use of the unlicensed spectrum in addition to the LTE licensed spectrum will be an important complement to meet the ultra-high capacity that will be needed for 4G and beyond.

As we have seen in the previous points, the idea of heterogeneous networks is very common in LTE, and we are going towards an environment full of small cells that help us to reach very high capacities. Following this tendency, LTE-U is a technology thought to be used in small cells, especially public indoor cells (e.g. in shopping centres) or outdoor hotspots (e.g. in business districts).

Until today, Wi-Fi has been the most popular choice for radio access in the unlicensed space. However, studies have highlighted that LTE-A technology has significant gains over Wi-Fi when operating in this band. The main advantages for LTE-U over Wi-Fi as an access technology are:

- Better spectrum efficiency and coverage due to more advanced radio features such as robust FEC (Forward Error Correction), hybrid ARQ (Automatic Repeat request) or interference coordination.
- The same RAN can provide LTE data access in licensed and unlicensed spectrum.
- A simplified network management and tracking of KPIs (Key Performance Indicators) through a single RAN can be achieved.
- Improved network management and load balancing through tighter integration.
- Instead of continue pursuing LTE – Wi-Fi interworking, LTE-U is well integrated to the existing operator network, thus solving all authentication, Operations and Management and QoS issues.
- LTE ecosystem kinds of applications (machine-to-machine, device-to-device…) are exploitable in LTE-U.

Despite this advantages, the introduction of LTE-U brings a number of challenges to be addressed. As it works in unlicensed spectrum, LTE-U must support fair access of multiple LTE-U networks and multiple Wi-Fi networks. When demand exceeds capacity, each network should be able to access an equal share. When a particular network’s traffic demand is less than the spectral capacity of an equal share, that network should allow other networks to access the unused capacity. This will require LTE-U to adapt to the presence of other LTE-U and Wi-Fi networks, while Wi-Fi uses its current mechanisms. Therefore, issues such as coexistence with Wi-Fi systems operating in unlicensed spectrum, unpredictable interference to LTE from other technologies and coexistence among cells from the same or different operators need to be resolved.

2.3.1. Band definitions for LTE-U

Although it was agreed that the core technology should not be only for a concrete part of the spectrum, a clear focus is placed on unlicensed operation in the 5 GHz band. Fully harmonized global regulations do not exist in the case of this band. However, all major markets offer more than 300 MHz available in this band.

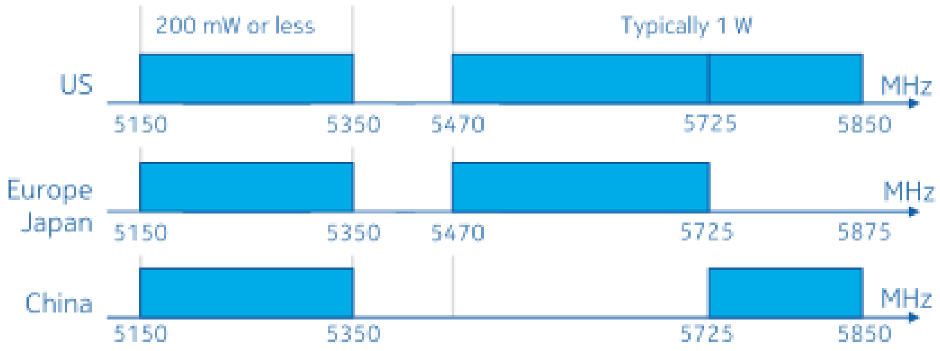


Figure 1. Unlicensed spectrum availability in different regions.[16]

2.3.2. Channel selection in LTE-U

Channel selection is the mechanism used to decide the operating channel (i.e. the central frequency and its associated bandwidth) where a small cell sets up a LTE-U carrier.

The LTE eNodeB should select a cannel that does not have another network operating on it with a high interference level, but rather select a channel that is either free or only slightly loaded. Ideally, the mechanism needs to choose the less loaded channel, because the cleaner the channel is, the more throughput the system get.

If interference is found in the operating channel and there is another cleaner channel available, the transmission can be switched to the new channel using LTE procedures. This ensures that the interference is avoided between the small cell and its neighbouring devices (Wi-Fi devices, other LTE-U small cells and/or others). With these procedures, we ensure that our LTE-U small cell is a “good neighbour”.

For certain bands such as 5.25-5.35 GHz and 5.47-5.725 GHz, there are further specific requirements imposed on channel selection mechanisms to allow the coexistence with radar systems. ETSI mandates a Channel Availability Check to detect the presence of radar signals in the different channels. Therefore, only channels without radar signals detected are available and can be selected for operation.

As a system able to select the best channel and change it when it is necessary is needed, a ‘smart’ mechanism has to be selected. A mechanism able to always know the best option. That’s why learning-based mechanisms have been developed in this area, because they give the system the ability to learn from the past situations and create a behaviour that always chooses the best option[8].

2.3.3. Channel Access in LTE-U [1]

Channel Access is the mechanism used to decide actual transmissions on the selected channel. It can be used as time-domain coexistence mechanism to allow that multiple devices share the same operating channel using a time-division scheduling or it can be used as frequency-domain coexistence mechanism using a frequency-division scheduling.

In Fig. 2 we can observe the two Channel Access mechanisms described before. In frequency-domain coexistence we can see how the different devices select different carriers. In time-domain coexistence, SC1 and SC2 have selected the same channel and they use a time-division scheduling, so they don't transmit at the same time.

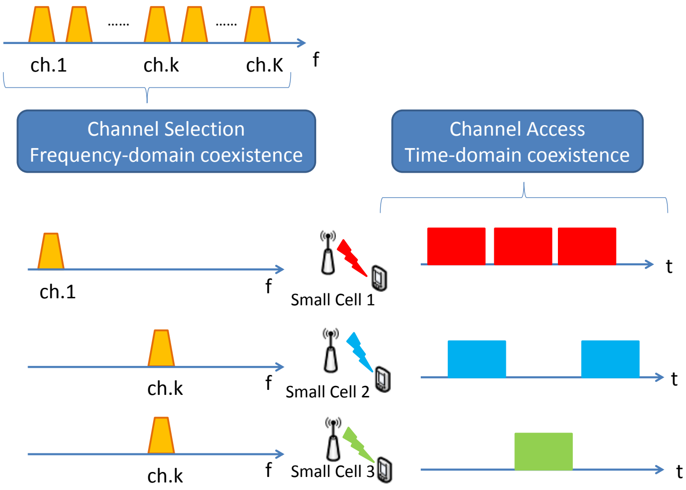


Figure 2. Illustration of the two coexistence mechanisms

In time-domain we can apply different strategies to ensure a fair coexistence. In some markets regulation requires the support of a Listen-Before-Talk scheme. For that purpose, a small cell using an LTE-U carrier will only transmit if it senses the channel as free (the received power in this channel is below a given threshold) during the Clear Channel Assessment time. Then, transmission will be done during a maximum time of 10 ms followed by an idle period, after which the CCA will be executed again.

2.4. Artificial intelligence and learning algorithms[17][18]

Artificial intelligence is the intelligence exhibited by machines or software, and it is also the name of the field of study which studies how to create computers and software that are capable of intelligent behaviour. In few words, this field is the study and design of systems that perceive its environment and takes actions according to it.

AI study is an interdisciplinary and vast field, which involves computer science, mathematics and specific fields like psychology or linguistics depending on the case it needs to be used, e.g. linguistics are needed if we want to create a speech recognition system.

The number of AI applications is infinite. We can find artificial intelligence from computer science or industry (robots that perform dangerous or repetitive jobs) to toys and computer games or hospitals.

In mobile communications, the increased network complexity and the need to provision a good quality of experience (QoE) to the end-users have made the traditional processes such as the network dimensioning, resource provisioning and the integration of new network elements become much more dynamic and complex. These networks should be also adapted to a much wider variety of scenarios and use cases. For these reasons, there

[8] Learning and learning algorithms are explained in the section 2.4.

is the need to include AI in this area, exploiting cognitive capabilities that embrace knowledge and intelligence.

Following this idea, mobile communications systems should become much more self-autonomous and be able to smartly process all the possible available inputs to eventually make the proper decisions. In this way, the networks become more efficient, dynamic and adaptive. This leads to network management mechanisms based on AI concepts that support the decision making and therefore leading to more efficient tactical and strategic decisions.

In our case, we want our small cells to learn from their own experience in order to make the most appropriate decisions in each case, so we need a learning algorithm. Machine learning is the field of AI needed when we want our system to improve through experience automatically. In the channel selection case, we want our algorithm to “discover” which is the best channel using its experience from previous situations.

2.4.1. Machine learning[18][19]

Machine learning explores the construction and study of algorithms that can learn from and make predictions on data. In other words, its goal is to build computer systems that can adapt and learn from their experience. Such algorithms operate by building a model from example inputs in order to make data-driven predictions or decisions, rather than following strictly static program instructions.

We can classify machine learning tasks in three categories:

- **Supervised learning:** The system is taught by showing in example inputs and their desired output. Its goal is to find a general rule that maps inputs to outputs. So it is not valid for problems where the desired behaviour is not known. It is used in classification and prediction models.
- **Unsupervised learning:** No labels are given to the learning algorithm, leaving it on its own to find structure in its input. It can be useful to find hidden patterns in data.
- **Reinforcement learning:** It consists in learning how to map situations to actions so as to maximise a scalar reward. In this case the learning is achieved through the interaction with the environment, so that the learner discovers which actions yield the most reward by trying them. There exist different categories of RL mechanisms: dynamic programming, Monte Carlo methods and Temporal-Difference learning. In dynamic programming a perfect model of the environment is needed, while in Monte Carlo methods no complete knowledge of the environment is needed, but a policy should be followed during a number of steps and only at the end of these steps the reward is obtained to improve the policy. Conversely, TD learning methods combine the benefits of the other two as they do not need a model of the environment dynamics and they are able to update the decision policy without waiting for the final outcome. In particular, TD learning methods adjust the estimated value of a state based on the immediate reward obtained after an action is made.

2.4.2. Q-learning

Q-learning is a TD learning method. It can be used to find an optimal action-selection policy for any given Markov decision process[9].

[9] A Markov decision process is a discrete time stochastic control process. At each time step, the process is in some state, and the decision maker may choose any action available in the current state. The process responds at the next time step by giving a corresponding reward and randomly moving into a new state. [23]

The algorithm executes an action in a specific state and gets the reward (a numerical score). The goal is to maximise the total reward. This is done by learning which action is optimal for each state, i.e. the action that returns the highest reward.

2.4.2.1. Q-learning-based channel selection in LTE-U[1]

In the case of channel selection, exploiting learning from past experience seems a pertinent principle in the LTE-U context. Each access point may autonomously learn what channels are usually not being used by its neighbours and then tend to select such free channels. Thanks to this learning capability, the number of scanning procedures that a base station needs to do to look for the cleanest channel can be reduced to a minimum. In addition, including adaptability to the learning-based decision-making process will provide robustness to the solution and the capability to react to changes in the scenario.

In our case, we will work in a Q-learning solution. The idea in this case is that each small cell progressively learns and selects the channels that provide the best performance based on the previous experience.

In particular, each small cell *i* stores a value function *Q(i,k)* that measures the expected reward that can be achieved by using each channel *k* according to the past experience. Whenever a channel *k* has been used by the small cell *i*, the value function *Q(i,k)* is updated following the next expression:

$$Q(i,k) \leftarrow (1-\alpha_L)Q(i,k) + \alpha_L \cdot r(i,k)$$

(1) Q-learning update expression

In this expression $\alpha_L \in (0,1)$ is the learning rate and *r(i,k)* is the reward that has been obtained as a result of the current use of the channel *k*. Assuming that the target of the channel selection is to find a channel that maximises the total throughput, the reward function considered is given by:

$$r(i,k) = \frac{\overline{R(\imath,k)}}{R_{max}}$$

(2) Reward formula

$\overline{R(\imath,k)}$ is the average throughput that has been obtained by the *i*-th small cell in channel *k* as a result of the last selection of this channel.

R(i,k) is calculated in the following way:

$$R(i,k) = \sum_{n=1}^{N(i)} \frac{B}{N(i)} S(SINR_n(i,k)) \frac{1-\theta_{idle}}{M(i,k)}$$

(3) Throughput served by a cell i *using the channel* k

N(i) is the total number of users being served by the small cell, *B* is the bandwidth of the channel, $SINR_n(i,k)$ is the signal to noise and interference ratio observed by the *n*-th user when downlink data is transmitted on the *k*-th channel, θ_{idle} is the fraction of time associated with the idle periods imposed by the LBT strategy, *M(i,k)* is the number of small cells that are sharing in the time domain the *k*-th channel with the *i*-th small cell following the LBT strategy and *S(·)* is a generic function ranging between 0 and S_{max} that provides the spectral efficiency in bps/Hz as a function of the SINR depending on the characteristics of the technology.

In (2), $R_{max} = B \cdot S_{max} \cdot (1 - \theta_{idle})$ is a normalization factor.

At initialization, *Q(i,k)* is set to an arbitrary value Q_0.

Based on the *Q(i,k)* value functions, the proposed Channel Selection decision-making for the small cell *i* follows the softmax policy in which channel *k* is chosen with probability:

$$\Pr(i,k) = \frac{e^{\frac{Q(i,k)}{\tau(i)}}}{\sum_{k'=1}^{K} e^{\frac{Q(i,k')}{\tau(i)}}}$$

(4) Channel selection probability

In this expression, *K* is the number of available channels and *τ(i)* is a positive parameter called 'temperature'. High temperature causes the different channels to be all nearly equiprobable. Low temperature causes a greater difference in selection probability for channels that differ in their *Q(i,k)* value estimates, and the higher the value of *Q(i,k)*, the higher the probability of selecting channel *k*. Softmax decision exploits what the system already knows in order to obtain reward, but it also explores to make better actions in the future. A cooling function is considered to reduce the temperature as the number of channel selections made by the small cell *i* increases, so that the amount of exploration will be progressively decreased as the small cell has learnt the best solutions. The specified cooling function is the next one:

$$\tau(i) = \frac{\tau_0}{\log_2(1 + n(i))}$$

(5) Cooling function

In this equation, τ_0 is the initial temperature and *n(i)* is the number of channel selections that have been already done by the *i*-th small cell.

2.5. Simulation scenario and code

2.5.1. Simulation scenario

The simulation code on which we worked propose an indoor scenario for LTE-U coexistence evaluation. It consists of a single floor building where two operators have 4 small cells installed. Their distribution is shown in the following image:

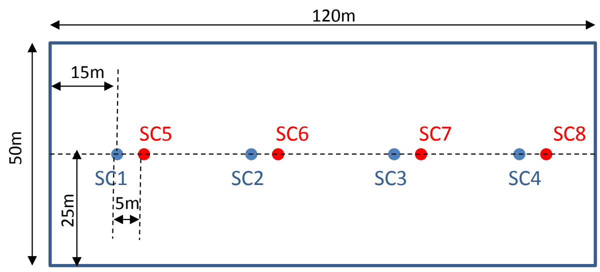


Figure 3. Distribution of the small cells in the simulator scenario.

In the figure, SC1 to SC 4 are owned by Operator 1 and SC5 to SC8, by Operator 2.

The small cells that have the same owner are equally spaced and all of them are centred along the shorter dimension of the building.

The height of the SC is 6m, while the antenna height of the mobile terminals is 1.5m.

A total of 10 users per operator are randomly distributed inside the building. Each user is associated to the small cell of its own operator that provides the highest received power.

The SC-to-terminal and the SC-to-SC path loss and shadowing are computed using the ITU InH model that can be found in [20]. The carrier frequency is 5 GHz and the channel bandwidth B=20 MHz. The transmit power in one LTE-U carrier is 15 dBm. Omnidirectional antenna patterns are assumed with a total antenna gain plus connector loss of 5 dB. The terminal noise figure is 9 dB. The spectrum efficiency form is obtained from Annex A.1 in [21] with S_{max}=4.4 b/s/Hz.

The threshold used in the CCA of the LBT to decide if a channel is sensed as free or not is -70 dBm/MHz. With this threshold and the considered propagation model it turns out that in the layout shown in Fig. 3 only SC3 and SC6 are able to detect the transmissions of all the other small cells. SC1 is not able to detect SC4, 7 and 8; SC2 does not detect SC8; SC4 does not detect SC1 and 5; SC5 does not detect SC4 and 8; SC7 does not detect SC1 and SC8 does not detect SC 1, 2 and 5.

The parameter θ_{idle} is set to 0.05 and the parameters related to Q-learning algorithm are α_L=0.1, τ_0=0.1 and Q_0=0.5.

Simulation time is measured in "time steps" (ts). It is considered that all the small cells are continuously generating geometrically-distributed activity periods with average 150 ts in which they require the activation of a LTE-U carrier to transmit data to their users.

2.5.2. Simulator code

We were provided with a Matlab code that simulates an LTE-U network with the conditions described in the previous point. This code is useful to test the small cells' performance in several cases, such as using a fixed channel or using a Q-learning channel selection mechanism.

In this section we are going to write a summary of this code.

CODE SUMMARY:

- Declaration of constants (5)[10]
- Setting of environment parameters: (34)
    - Size of the building (x and y)
    - Height of the access points
    - Height of the UEs
    - Frequency and BW
    - Propagation model constants (46)
    - Number of users per operator (56)
    - UE parameters (noise figure, noise power…)
    - AP parameters (66)
    - LBT parameters
    - Throughput computation parameters ($SINR_{min}$, S_{max}…) (80)
    - Number of APs per operator and distance between them
- Activity parameters: (94)
    - Creation of vectors activity_time and inactivity_time, which have one place for each AP.
    - Assignation of the values of activity and inactivity time for each operator.
- Channels:
    - Setting of the number of channels and the channel assignation in case of fixed allocation method is chosen.
- Operator algorithms: setting of what algorithm uses each operator (Q-learning, fixed allocation or random). (105)
- Q-learning parameters:

[10] These numbers between parentheses are the number of the line where we can find this part of the code.

- Setting of α_L, τ_0, Q_0, the kind of cooling (no cooling, cooling time or cooling samples) and the kind of reward (individual, aggregated total or aggregated per operator).

- Setting of the threshold to decide a channel change (measured as a ratio between average throughput and maximum throughput) and the maximum number of ts below this threshold to decide a HO. (121)
- Setting of maximum time steps and number of experiments.
- Setting of the number of simulation loops in case starting with different temperature values was desired.
- Creation of the variables to show the statistics (global average bitrate, operator average bitrate…) (131)
- FOR (iterating through all τ_0 values): (172)
  - FOR (iterating for all the number of experiments)
    - Definition of variables (detect condition, distance between APs, shadowing losses between APs, Q values, channel selection probability…) (182)
    - Creation of temporal evolution statistics matrixes (activity evolution, assigned channel evolution, Q evolution…), it is going to store data from each time step to this matrixes. (218)
    - Initialisation of the APs: setting their position and their operator. (229)
    - Check detection conditions between APs if they are operating in the same frequency. It calculates the attenuation and the shadowing between each pair of APs. With these values it calculates the received power and decides if they are detecting each other of not comparing the Rx power to a threshold. (257)
    - Initialisation of the UEs. It allocates each UE in a random position. It also calculates the shadowing and the losses between each UE and each AP. With these values, it connects each UE with the AP of its operator that provides more power in that place. (309)
    - Initialisation of the activity. It sets all the sessions as 'off' and calculates the probabilities of start and finish a session. (367)
    - Initialisation of the channels. It sets the algorithm that each channel uses and the initial value of Q. (378)
    - Simulation process. (395)
      - It evaluates the optimum combination. It creates all the possible channel combinations, evaluates all the combinations and chooses the one that has the best total theoretical bitrate as the optimum one.
      - Simulation starts. (503)
      - FOR (all the time steps)
        - Measure reward and statistics for all the APs that are active in time step t. (525)
        - FOR (all APs)
          - Fill vectors of temporal evolution.
          - FOR each channel, fill Q evolution vector, calculate temperature and selection probability.
          - If session is 'on', FOR (all the UEs the AP is serving): calculate the interference of each AF, capacity, APs that share channel and detect each other and update statistic variables. (547)

            - It checks the activity and the channel that is going to be used for each AP in the next ts. (607)
            - FOR (all APs)
                - If session is 'on': (610)
                    - Calculate if it is has to finish in this ts.
                    - If it finishes: calculate reward, update Q and release the channel.
                    - If it does not finish: check if it has to change its current channel, and do it if necessary. To select the channel it does it randomly, but following the channel selection probabilities, that calculates also here using the Q values (in case of Q-learning channel selection, in case of fixed channels it does not change and in case of random selection it selects the channel randomly from an equiprobable set of channels). (653)
                - If session is 'off': (732)
                    - Calculate if it has to start a session, and do it if necessary. To start a session it selects the channel in the same way as before.
        - When it has completed this for (for all the time steps) the experiment has been completed.
    - Calculates all the global statistics of the experiment (average bitrate per experiment, average bitrate per operator per experiment…). And here finishes the experiment. When all the experiments have finished, the simulation is done. (796)
  - Computes global statistics for all the experiments of a simulation. (812)
- Plot the results (833)
  - Plot 1: Q(channel)/t for each AP.
  - Plot 2: Channel selection probability/t for each AP.
  - Plot 3: Bitrate evolution for each AP.
  - Plot 4: Activity and assigned channel evolution for each AP.

3. Project development and results:

In this section we are going to talk about the different experiments realised in this project. As the project did not consist in only one experiment, the results are also included here in order to make the comprehension clearer.

3.1. Study of the impact of the initial value of Q on the convergence time of the system

The finality of this study is find the optimum initial value of Q (Q_0) that makes the channel selection probabilities converge as fast as possible.

To perform the experiment several simulations with different Q_0 values have been done, starting at 0.0 and finishing at 1.0 with a step of 0.1 or, in some special occasions, 0.05 or 0.2. The followed criteria to calculate the convergence time is that we considered convergence when the channel selection probability of one of the four channels that can be used was 99% or higher during 20000 ts in each base station.

Before starting the simulations, we needed to change a couple of parameters from the original code:

3.1.1. Environment

The simulation environment was the following:

- 2 operators.
- 4 base stations per operator (8 in total).
- Both operators use Q-learning mechanism.
- 4 available channels.
- 20 users per operator. We had to increase the number of UEs per operator from 10 to 20 to make it more difficult for the APs and increase deliberately the convergence time because with the default values it converged too fast with all the different values of Q_0.
- 20 experiments per simulation.
- 500k time-steps-long experiments. We also had to change the temporal length of the experiments from 1Mts to 500kts, because the APs normally reached the convergence between 0 and 50kts, so it was not necessary to perform 1M of time steps.
- The remaining parameters had their default values.

3.1.2. Simulation results

We made four complete simulation, i.e. with the same random number seed we tried different Q_0 values between 0.0 and 1.0. The first three simulations were used to find the tendency of convergence times in function of Q_0 and the last one to know if the results were similar if we changed the reward calculation mechanism from ‘global’ (default, it takes in account all the base stations from all the operators) to ‘per operator’ (that calculates the reward only using data from same operator’s base stations).

We stored the values of the convergence time of each base station in each simulation (as the criteria was the time that the probability was over 99% only one result per base station is possible) and calculated the mean of these values. In several cases, we found that no channel reached a selection probability of 99% (or it arrived considerably later than the other APs). In consequence, we calculated an ‘alternative mean’, called *x’*, only with the values that did not present problems.

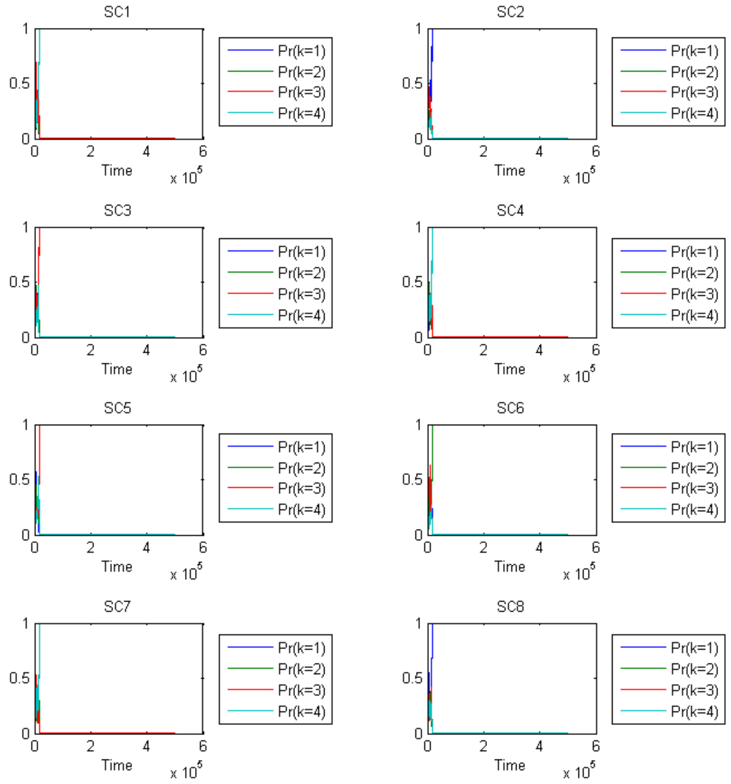


Figure 4. Example of the channel selection probability plots.

These previous plots are an example of the plots we obtained in each simulation when it calculated the mean channel selection probabilities from each AP during the 500kts of the experiments. With the same data used to create the plots we calculated the convergence time with a Matlab script that we had to create specially for this case.

In the next tables we can find the values obtained in each group of simulations:

First group of simulations										
	#Base station									
Q0	1	2	3	4	5	6	7	8	x	x'
0	1439	1717	1062	1853	3341	3350	1904	3033	2212.375	2212.375
0.1	2318	3768	1776	3111	22798	3984	2499	3493	5468.375	2992.714
0.2	3219	5282	22095	4689	4994	5172	2842	3582	6484.375	4254.286
0.25	3094	4137	4539	3141	3633	4444	2268	3485	3592.625	3592.625
0.3	4340	5109	4982	3730	5754	5273	3648	3523	4544.875	4544.875
0.35	4529	4137	4521	5310	4417	7545	4160	3524	4767.875	4767.875
0.4	10951	10939	12139	11856	12380	12764	12415	9929	11671.63	11671.63
0.5	21574	22797	20368	inf	20234	21678	21473	21347	-	21353
0.7	27017	26334	26800	26307	25936	inf	26514	26474	-	26483.14
0.9	31608	31286	29163	inf	29109	30605	31757	30539	-	30581
1	32849	34343	34610	32375	32297	32273	32540	92728	40501.88	33041

Table 2. Results of the first group of simulations.

Second group of simulations										
	#Base station									
Q0	1	2	3	4	5	6	7	8	x	x'
0	1072	1874	1971	2021	1450	692	2872	854	1600.75	1600.75
0.1	1558	1936	2059	3453	1621	907	3166	1218	1989.75	1989.75
0.2	3392	2270	2846	3686	2006	1550	3343	2275	2671	2671
0.25	3957	3143	2985	4037	3220	3361	3568	4489	3595	3595
0.3	5438	5837	5312	6395	4947	319610	4726	6858	44890.38	5644.714
0.4	5034	5837	5312	5783	4714	6341	6643	5556	5652.5	5652.5
0.5	16315	14779	14293	14585	13885	15561	15482	14031	14866.38	14866.38
0.6	14448	14547	14065	13955	13817	inf	14405	14049	-	14183.71
0.7	18411	18996	18311	20354	19300	19467	18406	18701	18993.25	18993.25
0.8	26411	26637	26820	27072	25843	26288	26609	26197	26484.63	26484.63
0.9	65317	inf	57068	81637	54431	60442	54438	56600	-	61419
1	19993	21804	19682	20354	20672	20810	19883	inf	-	20456.86

Table 3. Results of the second group of simulations.

Third group of simulations										
	#Base station									
Q0	1	2	3	4	5	6	7	8	x	x'
0	1072	1874	1971	2021	1450	692	2872	854	1600.75	1600.75
0.1	1558	1936	2059	3453	1621	907	3166	1218	1989.75	1989.75
0.2	3392	2270	2846	3686	2006	1550	3343	2275	2671	2671
0.3	5438	5837	5312	6395	4947	319610	4726	6858	44890.38	5644.714
0.4	5034	5837	5312	5783	4714	6341	6643	5556	5652.5	5652.5
0.5	16315	14779	14293	14585	13885	15561	15482	14031	14866.38	14866.38
0.6	14448	14547	14065	13955	13817	inf	14405	14049	-	14183.71
0.7	18411	18996	18311	20354	19300	19467	18406	18701	18993.25	18993.25
0.8	26411	26637	26820	27072	25843	26288	26609	26197	26484.63	26484.63
0.9	65317	inf	57068	81637	54431	60442	54438	56600	-	61419
1	19993	21804	19682	20354	20672	20810	19883	21497	20586.88	20586.88

Table 4. Results of the third group of simulations.

As we can see in the previous tables, some numbers are highlighted. These numbers correspond to the maximum value of the convergence time in each simulation. If it is highlighted in green it means that is close to the other values, and in red that it's far from the other values. When 'inf' is written in the table it means that that base station did not arrive to a convergence status during the experiment. At the right of the table we can find the mean *x* and the alternative mean explained before, *x'*.

In the next figure we can observe the evolution of *x'* depending on Q_0. We can see that they follow the same tendency (except for this peak at 0.9) that the second and the third have.

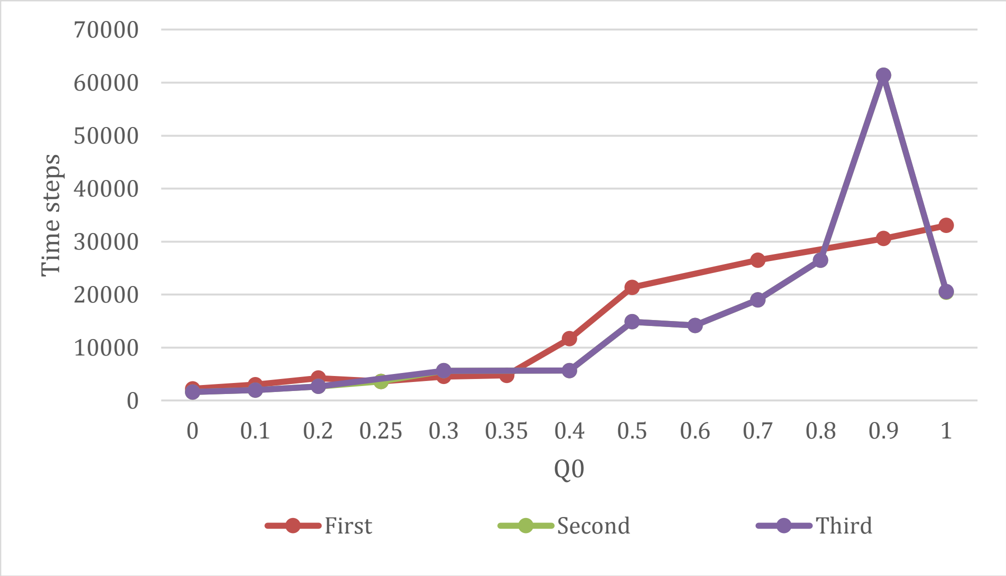


Figure 5. Evolution of the convergence time in function of Q_0.

As we told before, a fourth group of simulations was performed in order to know if the results changed if we changed the reward calculation method.

As we can see in the figure below, the fourth group of simulations (using a different reward calculation method) got the same tendency as the other three groups.

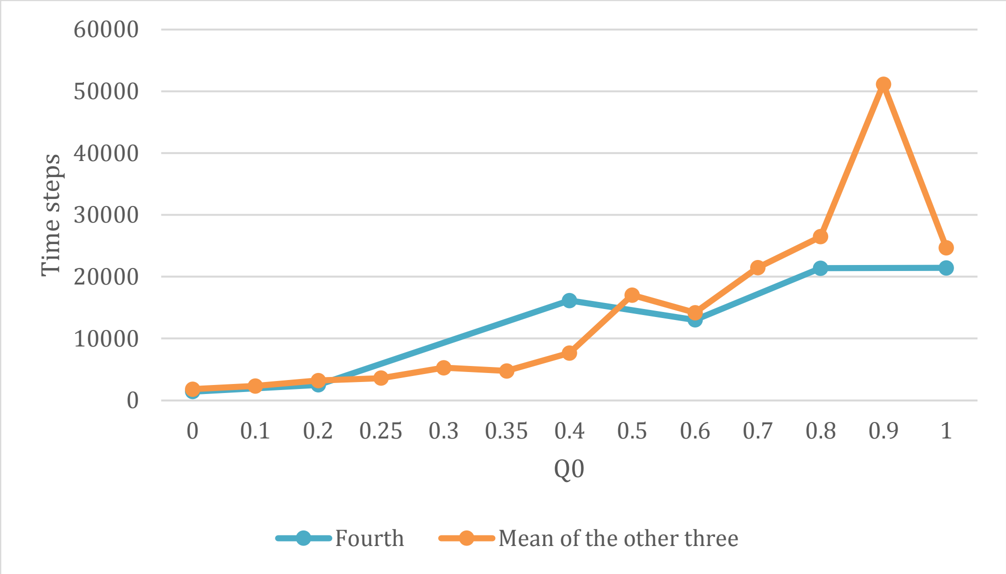


Figure 6. Comparison between the fourth group of simulations and the mean of the other three groups.

In both figures (Fig. 5 and Fig. 6) we can observe that the value that gives us the minimum convergence time is $Q_0 = 0$. In this way we can affirm that the optimum value for Q_0 is the 0, and with the system gets the minimum convergence time.

What we observed is that, the lower the Q_0 is, the lower the final Q value is. This means that the Q value converges to lower values when Q_0 is low.

In the next figure we can see the Q evolution when Q_0 is 0.1:

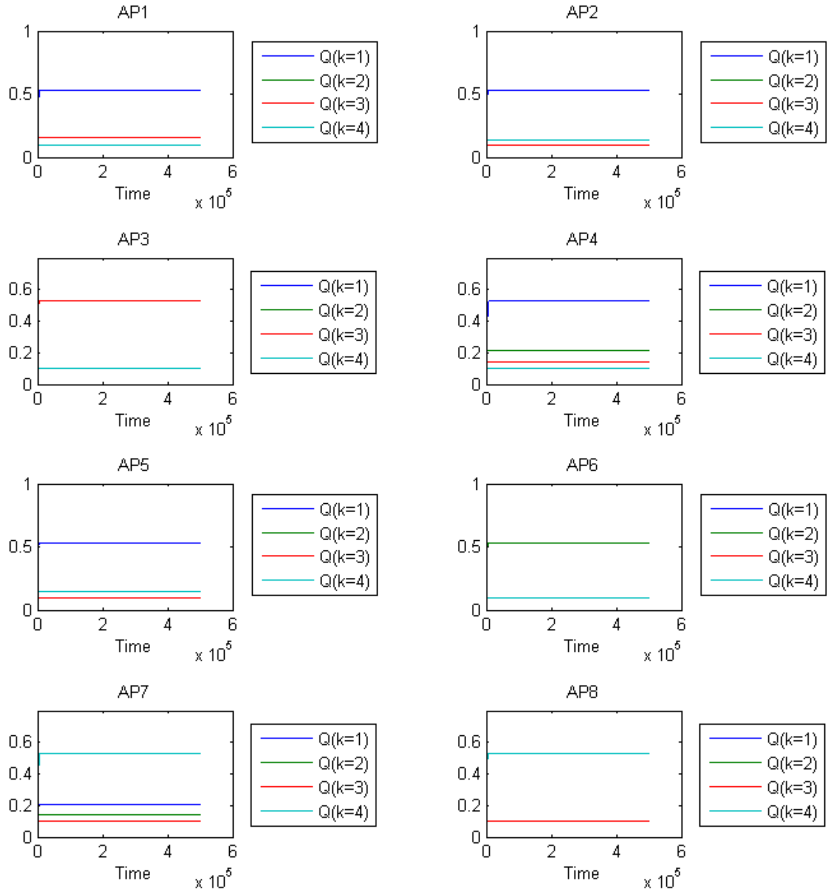


Figure 7. Q evolution when Q_0 = 0.1.

We observe that all Q values converge to values near 0.5. However, if we use a higher Q_0 we get the following results:

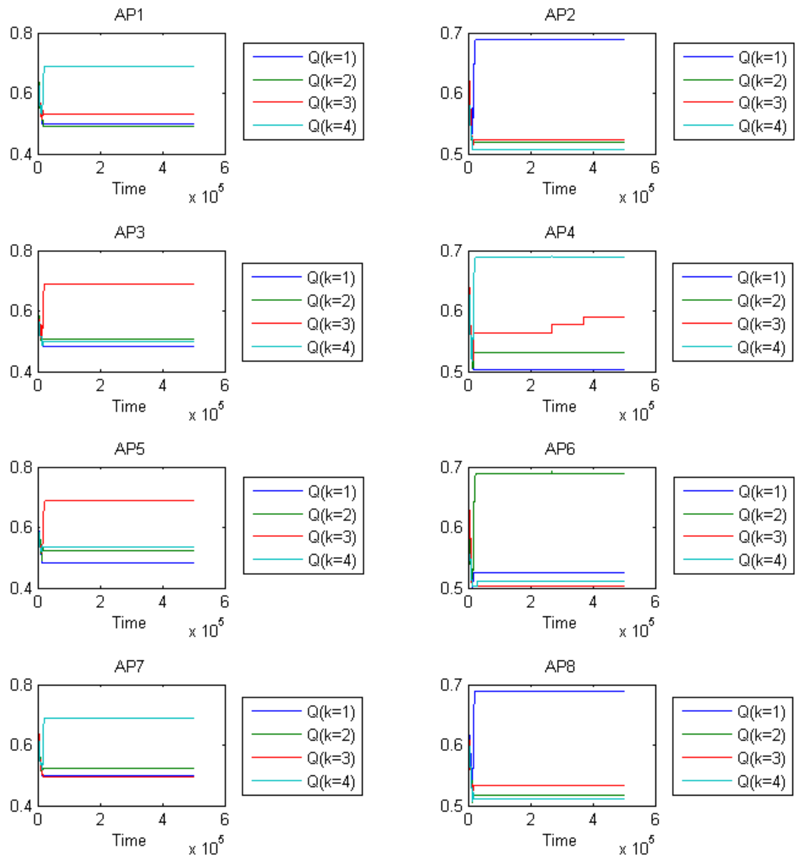


Figure 8. Q evolution when $Q_0 = 0.7$.

As we can see in the previous figure, Q converge to values near 0.7.

Even though this fact a priori can seem a bad, it is not bad because Q does not affect to the performance of the base stations. As Q is only used to compare channels between them, the important result is the difference between the Qs of the different channels, not their numerical value.

To sum up, we found that the value of Q_0 that provides the system with a lower convergence time is 0.

3.2. Study of the impact of the number of channels on the convergence time of the system

In this study we wanted to discover how much the convergence time changes in function of the number of available channels.

3.2.1. Environment

The following environment was set:

- $Q_0 = 0$, as we found this value as the optimal one.
- 4 or 8 channels, depending on the simulation.
- 2 operators.
- 4 base stations per operator.

- 15 users per operator.
- Reward calculated per operator.
- 10 experiments in each simulation. We needed to lower this because with 20 experiments per simulation the simulations were too long in time (4-5 hours) in the case of 8 available channels.
- Experiment time: 500k ts.
- All the remaining parameters were set as default.

3.2.2. Simulation results

After 4 simulations with 4 available channels and 4 more with 8 available channels, we got these results using the same Matlab script that we used in chapter 3.1:

	Sim. ID[11]	#Base Station									
		1	2	3	4	5	6	7	8	x	x'
4 ch.	700	982	1808	inf	1557	1076	2252	1064	2126	-	1552.143
	742	2959	1639	2646	915	1727	1728	1081	1241	1742	1742
	234	1694	2388	1834	1509	1088	3089	2692	2932	2153.25	2153.25
	47	1436	3684	1302	inf	998	1579	905	1212	-	1588

	Sim. ID	#Base Station									
		1	2	3	4	5	6	7	8	x	x'
8 ch.	700	982	1633	inf	1381	1000	2301	879	1405	-	1368.714
	742	1768	1735	1029	915	2148	2173	1001	1241	1501.25	1501.25
	234	946	1392	2292	678	1671	1374	1691	2024	1508.5	1508.5
	47	1436	1639	933	inf	1858	1579	905	1012	-	1337.429

Mean	
4 ch.	8 ch.
1758.8	1429.0

Table 5. Results of the convergence time calculation with 4 or 8 available channels.

In these tables we can observe the results of 4 different simulations in each case and also the mean of the results obtained in both cases. As in the previous chapter, we had to create the variable *x'* in order to get a mean value when one of the values in the table was 'inf', that means that the base station did not reach the convergence within the time of the experiment.

In this case, the reason for this APs to not reach the convergence is that they did not have any connected UE, so they were 'off'. We can know this because the Q of all channels is always 0 (as we can see in the figure below), it means that the Q is never updated because the algorithm only updates Q when the AP is 'on', and the APs are 'on' only when they have at least one UE connected to them.

[11] Sim. ID: Simulation Identification. Is the number that we used to identify each simulation, which corresponds to the seed of random numbers.

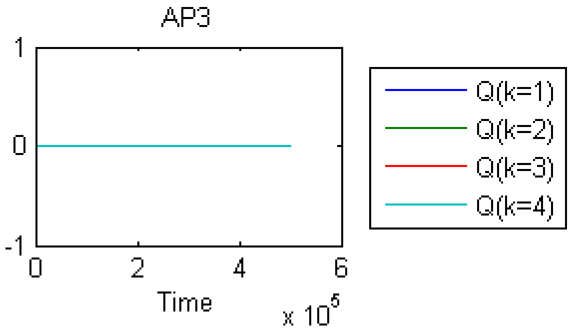


Figure 9. Q evolution of AP3 in simulation 700 with 4 available channels.

As we can see in the previous tables, increasing the number of available channels makes the convergence time decrease. This is due to the less number of 'conflicts' (neighbour base stations using the same channel) that are produced when the base stations can choose between 8 channels instead of 4, in this case. It is true that the time of convergence decreases, but it does not do it dramatically. In this case, we doubled the number of channels and the convergence time decreased an 18.8%.

3.3. A new mechanism to improve the convergence time of the system: Semi-fixed Q-learning

Finding the optimum value of Q_0 we reached a minimum value in terms of convergence time of our system. Anyway, more improvements could be done in order to lower even more this time.

The solution we propose is a semi-fixed Q-learning. With the word "semi-fixed" we mean that we set a group of channels which can be used by the access point, so the AP can only choose between a set of channels instead of all the available channels in the spectrum.

The idea came when we were analysing the environment in which the access points were working. We realized that neighbour cells wouldn't use the same channel very often, so we thought that we could reduce the work of our access points if each of them knew the channel that their neighbour was using.

However, we wanted a system that could operate independently in each AP, so we had to think about a mechanism that could be implemented in every access point without the need of any shared infrastructure between access points.

The easiest solution that we came up with was giving the APs only two channels (over four) to choose. Thus they can reduce dramatically the number of operations.

In the next figure (the same as Fig. 3) we can see the distribution of the access points. Blue ones are from Operator 1 and red ones from Operator 2. As we can see, SC1 and SC5 e.g. are not supposed to use the same channel, so we can set to one of them a pair of channels and to the other we can set the other pair and they will never collide.

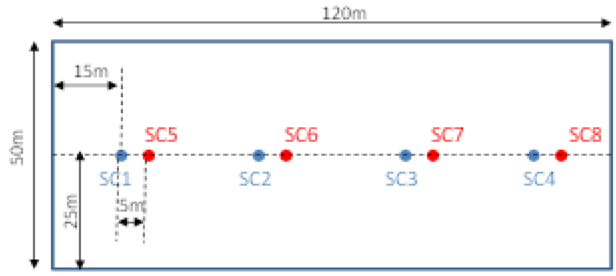


Figure 10. Distribution of the APs.

For example, with this layout we can set channels 1 and 2 to SC1, SC3, SC6 and SC8; and channels 3 and 4 to SC2, SC4, SC5 and SC7.

From the point of view of the installation, as the installer knows where the access points are located, the group of channels that each AP can choose can be set when the installation is carried on. In case of another operator wanted do install their infrastructure in the same place, they can check the channels used by the previous operators' access points and configure their own APs with the best selection of channels in that case.

3.3.1. Coding

To check if this idea could make our system to decrease its convergence time we had to program the procedure in the piece of Matlab code we used since the beginning. We studied what was the best way to make the APs only take in account the channels we wanted.

The easiest solution we found was modifying the Q values of the channels that we didn't want to use. We created a matrix called 'canals_assignats' that had the information about the channels that each AP was allowed to use.

The reason for this modification is that we needed to "eliminate" two channels in every AP, but as they were always different channels, we could not reduce the number of channels to two.

So, analysing the mathematical equations for the channel selection probability we found that reducing the Q was a good way to make the channels have a probability of 0 (or very close to 0).

We knew that:

$$\Pr(k) = \frac{e^{\frac{Q(k)}{\tau}}}{\sum_{k'} e^{\frac{Q(k')}{\tau}}}$$

(6) Channel selection probability formula.

In this case we can do the following approach:

$$\sum_{k'} e^{\frac{Q(k')}{\tau}} = e^{\frac{Q(1)}{\tau}} + e^{\frac{Q(2)}{\tau}} + e^{\frac{Q(3)}{\tau}} + e^{\frac{Q(4)}{\tau}} \cong E(constant)$$

(7) Approach of the denominator of (6) to a constant.

As we wanted to make the Pr(*k*) of two channels be very low, we needed to make the numerator of the fraction very low (ideally 0), so the previous sum would be only the sum of the two channels that are allowed. We could approach this sum to a constant in the cases we were not calculating the Pr(k) of any of the allowed channels.

Carrying on with the first equation using now this approach:

$$\Pr(k) = \frac{e^{\frac{Q(k)}{\tau}}}{E} \cong 0 \rightarrow e^{\frac{Q(k)}{\tau}} \cong 0 \rightarrow Q(k)\ \downarrow\downarrow\downarrow$$

(8) Deduction using (6) and (7)

With this result we decided that the best way to make the channel selection probability be 0 was modifying the Q of undesired channels with a very low value in each ts, so we set these values to '-10'. In this way we increase a bit the number of operations that need to be done in each ts instead of reducing them.

3.3.2. Environment

To check if this mechanism worked, we used the following environment:

- 20 experiments in each simulation.
- 20 users per operator.
- 2 operators.
- 4 available channels.
- $Q_0 = 0$

The channel distribution in the simulation in which we used the semi-fixed Q-learning was the following:

- AP1, AP3, AP6 and AP8 used channels 1 and 2
- AP2, AP4, AP5 and AP7 used channels 3 and 4.

All the other input values were set as default.

3.3.3. Simulation results

From these two simulations we could gather that the new algorithm was more efficient than the old one. Lowering the number of time steps to get the convergence from 1836,8 time steps (the average in the Q-learning case) to 2 ts.

The main advantage of this algorithm is that it has a very low time of convergence. The other main advantage for this algorithm is that we avoid interference between APs when we do the installation, as we set the neighbouring APs to different channel sets.

To see how this algorithm responded to the introduction of an interfering signal, we tried to add an interference in the current channel of AP1. The reaction of the system was not the desired one, it started jumping from a channel to another and it ended only jumping between channels 1 and 2. We wanted it to change from 1 to 2 and remain in the second. After this experiment, we decided that forcing Q values was not a good idea, so we had to think of another mechanism that could solve this problem.

Logically, the drawback that it has is that we have not implemented a mechanism to change the set of channels when none of them are clean enough to have a good transmission (e.g. in an environment with lots of Wi-Fi routers occupying the spectrum), so if this mechanism found itself in a bad environment some of the APs would not work properly.

To sum up, we have to say that this new algorithm is very powerful in clean environments, but it needs improvement in order to work in noisy environments.

3.4. Semi-fixed Q-learning version 2

The main problem of the first version of semi-fixed Q-learning was that it was not robust to noisy environments. We decided to make a new algorithm that was simpler and resistant to interferences, because first algorithm was very slow (simulations of that algorithm took the double amount of time than ‘normal’ simulations, almost 2h).

The main difference between the two versions of the algorithms was that, in the first version, we were always setting the Q of the channels that we were not allowed to use and, in the second version, we never take into account the channels we are not allowed to use.

We set manually the available channels for each SC and modified all the code to take into account only that channels.

To make it robust against interferences we created a mechanism that changes the set of channels when we are using bad quality channels. The requirement for an AP to change its set is that it has to do 10 consecutive changes of channel due to bad quality. After this 10 handovers, it changes the set of channels in order to avoid the interference that it receives in the first set.

3.4.1. Environment

We set this environment to check the performance of our new algorithm:

- 20 experiments in each simulation.
- 20 users per operator.
- 2 operators.
- 4 available channels.
- $Q_0 = 0$
- Reward calculated per operator.

The default channel distribution in the simulation was the following:

- AP1, AP3, AP6 and AP8 used channels 1 and 2
- AP2, AP4, AP5 and AP7 used channels 3 and 4.

All the other input values were set as default.

3.4.2. Simulation results

What we first noticed after the simulation was that the convergence was different from the previous cases. In those cases one channel got better Q than the others and its selection probability increased to almost reach the 100%. However, in this case both available channels remained near a Q of 0.5. That made their selection probability to change very fast.

In the next figures we can observe this effect:

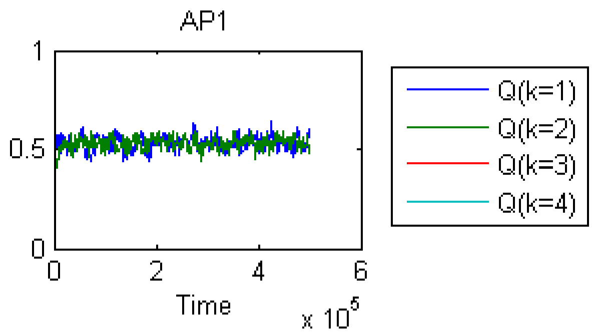


Figure 11. Evolution of Q in SC1 using Semi-fixed Q-learning v2

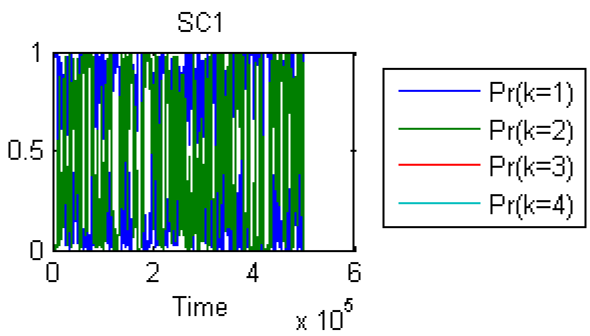


Figure 12. Evolution of channel selection probability in SC1 using Semi-fixed Q-learning v2.

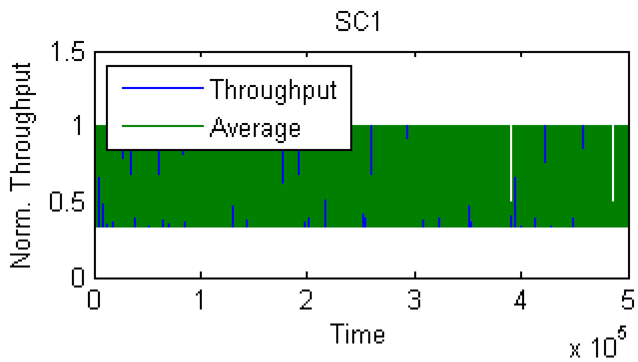


Figure 13. Evolution of the throughput in SC1 using Semi-fixed Q-learning v2.

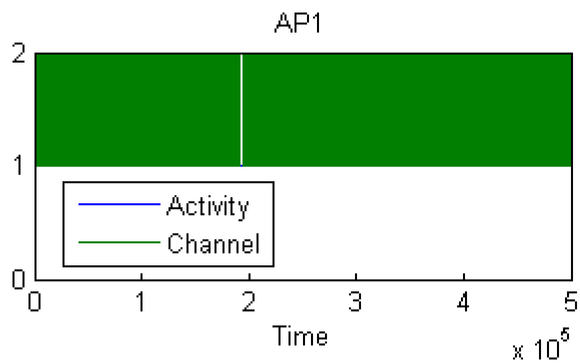


Figure 14. Evolution of current channel in SC1 using Semi-fixed Q-learning v2.

As we can see in the figures, in this case, the average selection probability of channels 1 and 2 would be approximately 50%. Due to this fact, every time the SC starts a session it can choose both channels with the same probability. This causes that in Fig. 14 the channel is always 'jumping' from 1 to 2, we can only see a 'green bar' because the average duration of a session is 150 ts and the length of the experiment is 500k ts. It is not bad because these changes are due to ending old sessions and starting new ones, not due to a bad quality of the channel. We can observe that the channel normalised throughput never goes below the threshold, which is 0.1 by default.

Due to this different behaviour, we need to create a new Matlab script to calculate the convergence time. We considered that the convergence was reached when the Q of any channel became more than 0.5.

The results after we calculated the convergence time applying the algorithm are the following:

#SC	1	2	3	4	5	6	7	8
Conv. time	3698	5884	9786	9716	9169	7680	5990	inf
x'	7417.571							

Table 6. Convergence time using Semi-fixed Q-learning v2.

We also needed to use the *x'* explained in previous experiments because no users were connected to SC8.

We can see that the mean convergence time is higher than in the previous chapters (we were talking about times of 1500 ts in the normal Q-learning cases (with Q_0=0) and 2 ts in the first version of Semi-fixed Q-learning). It is difficult to compare this case with the previous ones because the behaviour is completely different.

In order to study more this case and check its behaviour against interferences we included an interference in the zone of AP1 in channel 1 that started when *t* = 150k ts.

We can see the results in the next images:

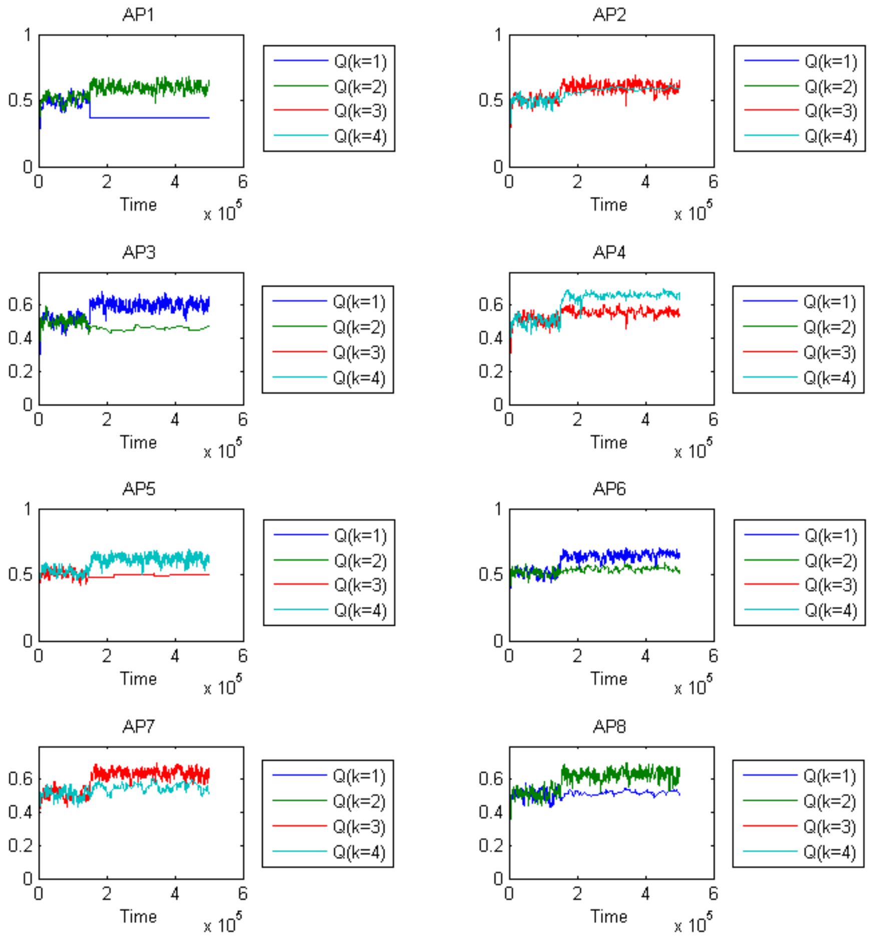


Figure 15. Q evolution when one interference has been included.

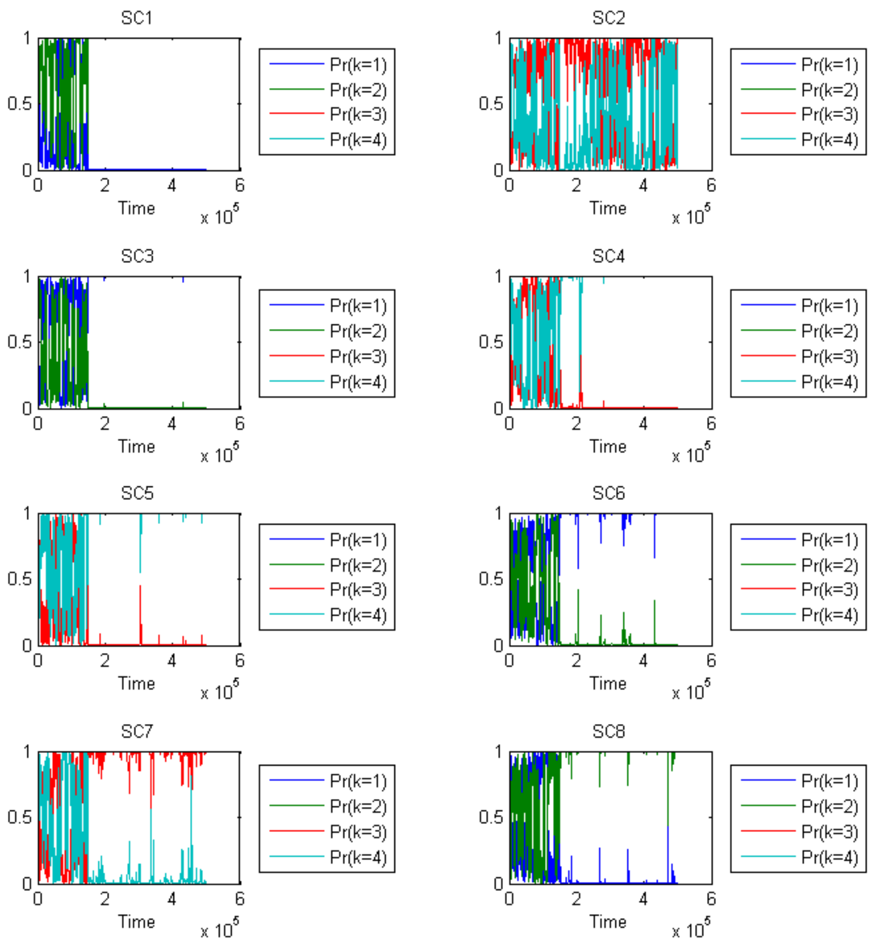


Figure 16. Channel selection probability evolution when one interference has been included.

We can observe that the system responds very fast to the interference, as AP1 increases the Q of channel 1 and it makes increase its selection probability. Therefore, all the neighbouring cells adapt to this new situation also fast.

We calculated the adaptation time in SC1 using the first Matlab script to calculate the convergence time, that calculated when a channel reached a 99% of selection probability and kept it, and the result was that SC1 adapted to this new situation (from having two free channels to giving channel 1 a 99% of probability) in 24 ts.

Afterwards, we checked if the mechanism to change the set of channels worked properly. To do this, we added another interference in channel 2 in the zone of SC1. With both channels with bad quality, SC1 had to change its set of available channels.

We can observe its performance in the next images:

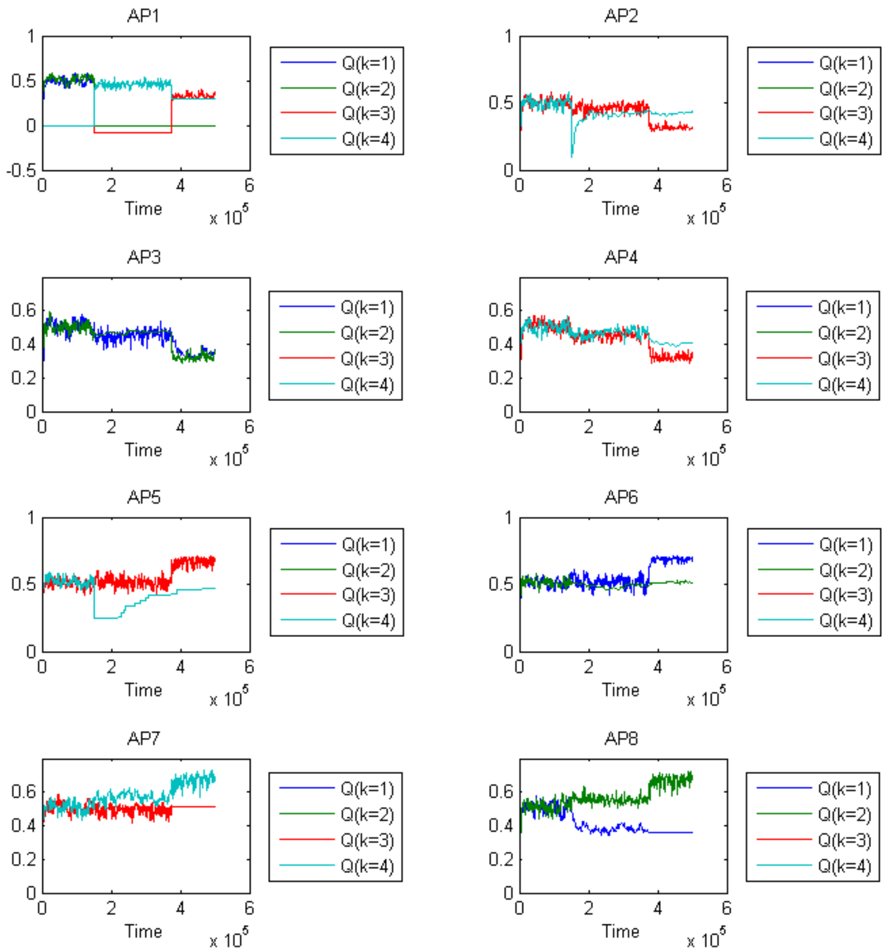


Figure 17. Q evolution with both default channels of SC1 interfered.

We can observe in SC1 how it changes from 1 and 2 to 4 (because 3 was used in that moment by SC5, which is near SC1) and afterwards to both 3 and 4.

We could check that the mechanism works properly.

To avoid the huge effect that the change has to other cells, we can set the reward as ‘aggregated total’. In this way we get better-looking results such as the ones shown in the next image:

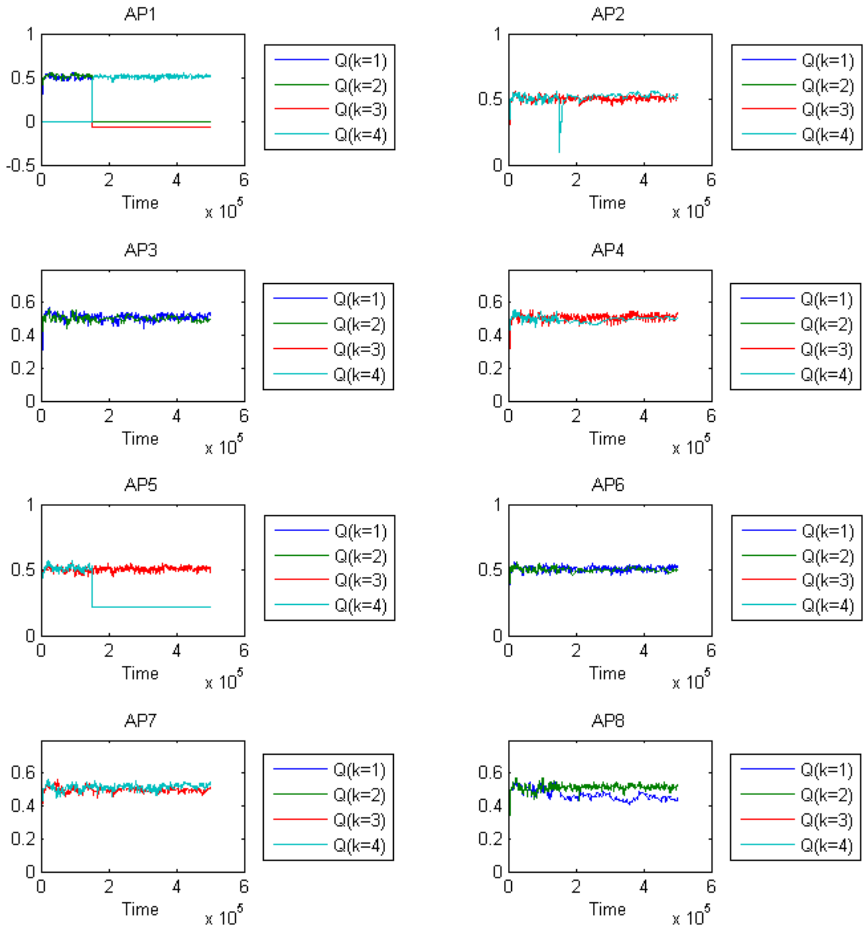


Figure 18. Q evolution with both default SP1 channels interfered and reward type 'total'.

Using this kind of reward, the other cells are less affected to the variations. This is due to the smaller variation that suffers the reward with this method. This smaller variation Q values also change less, so we get more stability.

We can see why the reward is more stable if it is calculated with the 'total' than with the 'per operator' method:

$$r_{TOTAL} = \frac{\frac{\sum_{i=1}^{A} \overline{R_{TOT}(\iota,k)}}{A}}{R_{TOT,max}}$$

(9) Formula to calculate r_{TOTAL}.

$$r_{OPERATOR} = \frac{\frac{\sum_{i=1}^{B} \overline{R_{TOT}(\iota,k)}}{B}}{R_{TOT,max}}$$

(10) Formula to calculate $r_{OPERATOR}$.

In the previous formulas, *A* is the total number of SCs and *B* is the number of SCs that the operator has. As always *A* is equal or bigger than *B*, a variation in the R_{TOT} of a SC has more impact (or the same in case that we have only one operator) when we are calculating the reward aggregated per operator.

In conclusion, this second version of the Semi-fixed Q-learning can resist to interferences and react very fast. In comparison with the other mechanisms, this mechanism normally "converge to two channels" instead of one, like the other mechanisms normally do.

4. Budget

As this thesis was not about building any prototype, we only need to take into account the software cost and the amount of hours dedicated.
The software cost is the following:

Software license	Cost
Matlab for academic use	500 €

Table 7. Total software cost.

In addition to Matlab, we used other software (like image edition software, data sheets, text processor or references organisation software), but this software was open-source, so it does not cost money.

This thesis took around 300 hours during all the semester, so their cost is the following:

Number of total hours	Cost/hour of a junior engineer	Total cost
300 h	10 €/h	3000 €

Table 8. Total hour cost.

As we can see, the total cost of this project is 3500€.

5. Conclusions and future development:

Nowadays, machine learning is one of the most studied fields of artificial intelligence. As the technology evolves fast, every day systems are more complex and the users demand more to their devices and technology, we arrived to the point that technology cannot work properly if it does not adapt to the environment or to the user. We can say that in some years artificial intelligence, especially machine learning, will be 'everywhere'.

Learning solutions in mobile communications will be one of the most important upgrades in the following generations, as they provide the devices with the ability to learn from experience and adapt to the environment trying to always choose the best option to maximise their performance.

These learning solutions are perfect for the channel selection algorithms in difficult environments such as LTE-U, where we have to deal with Wi-Fi signals and other access points that interfere our connections and we have to provide our device with a mechanism to avoid these kind of problems autonomously.

About our learning solution, in this project we have improved it finding and using its optimal Q_0 value and proved that the algorithm converges faster as more channels it can choose. Moreover, we proposed two enhancements to that algorithm: the first one got a very fast convergence but its reaction to interferences was poor; the second one overcame these drawbacks. Even though this second enhancement makes the base stations less 'free', they get a good response and a very fast adaptation to the interferences, and this 'loss of freedom' makes them avoid the interferences with the neighbouring small cells since the first moment. This variation of the algorithm also reduces the amount of operations that need to be done in each time step.

Finally, even though in this project we analysed and improved the channel selection algorithm, there is still a lot of work to do in this area:

- We could observe that the convergence time in function of Q_0 tendency was always the same. But, in order to enforce this result, a study with more simulations and different cases could be done to prove if the optimal Q_0 value is the same that was found in this project and/or if it changes in special cases.
- An organisation of the available LTE-U channels could be done. Some papers talk about using the same organisation as the Wi-Fi channels. So, analysing the environment and deciding how many channels can be allocated and their bandwidth and central frequencies could be an interesting study.
- The Semi-fixed Q-learning algorithm can be improved. It needs more testing, especially testing it in different environments and proving it against problems that can be found in the 'real life'.
- The initial algorithm could be tested in worst conditions, trying to create an environment more similar to the reality; see its performance and improve it whether it is necessary.
- The algorithm could also be improved using Carrier Aggregation.

Bibliography:

[1] O. Sallent, J. Pérez-Romero, R. Ferrús, and R. Agustí, “Learning-based Coexistence for LTE Operation in Unlicensed Bands,” IEEE ICC 2015 - Workshop on LTE in Unlicensed Bands: Potentials and Challenges, June, 2015.

[2] ETSI MCC, “Overview of 3GPP Release 8 V0.3.3.” 3GPP, September, 2014.

[3] C. Gessner, A. Roessler, and M. Kottkamp, “UMTS Long Term Evolution (LTE) Technology Introduction,” Application Note, Rohde & Schwarz, July, 2012.

[4] R. Agustí, F. Bernardo, F. Casadevall, R. Ferrús, J. Pérez, and O. Sallent, “1.3.1 LTE,” in *LTE: nuevas tendencias en comunicaciones móviles*, Fundación Vodafone España, 2010, pp. 46–48.

[5] M. Kottkamp, J. Rössler, and J. Schlienz, “LTE Release 9 Technology Introduction,” White Paper, Rohde & Schwarz, December, 2011.

[6] ETSI MCC, “Overview of 3GPP Release 9 V0.3.4.” 3GPP, September, 2014.

[7] ETSI MCC, “Overview of 3GPP Release 10 V0.2.1.” 3GPP, June, 2014.

[8] M. Kottkamp, A. Roessler, and J. Schlienz, “LTE-Advanced Technology Introduction,” White Paper, Rohde & Schwarz, August, 2012.

[9] R. Agustí, F. Bernardo, F. Casadevall, R. Ferrús, J. Pérez, and O. Sallent, “1.3.2 LTE-Advanced,” in *LTE: nuevas tendencias en comunicaciones móviles*, Fundación Vodafone España, 2010, pp. 48–50.

[10] ETSI MCC, “Overview of 3GPP Release 11 V0.2.0.” 3GPP, September, 2014.

[11] A. Roessler and M. Kottkamp, “LTE- Advanced (3GPP Rel.11) Technology Introduction,” White Paper, Rohde & Schwarz, July, 2013.

[12] ETSI MCC, “Overview of 3GPP Release 12 V0.1.4.” 3GPP, September, 2014.

[13] 4G Americas, “3GPP Release 12 : Understanding the Standards for HSPA + and LTE-Advanced Enhancements,” Executive Summary, February, 2015.

[14] LTE-U Forum, “Coexistence Study for LTE-U SDL.” LTE-U Technical Report, February, 2015.

[15] D. Flore, “3GPP & unlicensed spectrum,” 3GPP, IEEE 802 Interim Session, Atlanta, USA, January, 2015.

[16] Nokia Networks, “Nokia LTE for unlicensed spectrum,” White Paper, Nokia Corporation, 2014.

[17] Wikipedia, “Artificial intelligence.” [Online]. Available: https://en.wikipedia.org/wiki/Artificial_intelligence#Planning. [Accessed: 01-Jul-2015].

[18] J. Pérez-Romero, O. Sallent, R. Ferrús, and R. Agustí, “Artificial Intelligence as a Cornerstone for 5G Networks Management.” White Paper, Dept. of Signal Theory and Communications, UPC, July, 2014.

[19] Wikipedia, “Machine Learning.” [Online]. Available: https://en.wikipedia.org/wiki/Machine_learning. [Accessed: 01-Jul-2015].

[20] 3GPP TR 36.814 v9.0.0, “Further advancements for E-UTRA physical layer aspects.” March, 2010.

[21] 3GPP TR 36.942 v12.0.0, “Radio Frequency (RF) system scenarios.” September, 2014.

[22] RCR Wireless News, “LTE Drive Test Test Types and Parameters for LTE Driving Testing,” 2014. [Online]. Available: http://www.rcrwireless.com/20140509/wireless/lte-drive-test-test-types-parameters. [Accessed: 30-Jun-2015].

[23] L. Kallenberg, “Finite State and Action MDPs,” in *Handbook of Markov Decision Processes: Methods and Applications*, Kluwer, 2002, pp. 21–31.

Glossary

3G: 3rd Generation of mobile communications

3GPP: 3rd Generation Partnership Project

4G: 4th Generation of mobile communications

AI: Artificial Intelligence

AP: Access Point

ARQ: Automatic Repeat request

CA: Carrier Aggregation

CCA: Clear Channel Assessment

CDF: cumulative density function

CoMP: Coordinate Multi-Point operation

CRS: Cell-specific Reference Signal

DL: Downlink

eICIC: Enhanced Inter-Cell Interference Coordination

eNB: eNodeB

eNodeB: Evolved NodeB

E-PDCCH: Enhanced Physical Downlink Control Channel

E-UTRA: Evolved UTRA

E-UTRAN: Evolved UTRAN

FDD: Frequency Division Duplex

FEC: Forward Error Correction

feICIC: Further Enhanced Inter-Cell Interference Coordination

GNSS: Global Navigation Satellite Systems

GSM: Global System for Mobile communications

GW: Gateway

HeNB: Home eNB

HetNet: Heterogeneous Network

HO: Handover

ICIC: Inter-Cell Interference Coordination

IDC: In-Device Coexistence

IMT: International Mobile Telecommunications

KPI: Key Performance Indicator

LBT: Listen-Before-Talk

LTE: Long Term Evolution

LTE-A: LTE Advanced

LTE-U: LTE in Unlicensed bands

MBMS: Multimedia Broadcast/Multicast Services

MIMO: Multiple-input Multiple-output

NAICS: Network Assisted Interference Cancellation and Suppression

OFDMA: Orthogonal Frequency Division Multiple Access

ProSe: Proximity Services

QAM: Quadrature Amplitude Modulation

QoE: Quality of Experience

QoS: Quality of Service

RAT: Radio Access Technologies

RAN: Radio Access Network

RL: Reinforcement Learning

Rx: Reception/Receptor/Received

SC: Small Cell

SC-FDMA: Single Carrier Frequency Division Multiple Access

TA: Timing Advance

TD: Temporal Difference

TDD: Time Division Duplex

ts: time steps

Tx: Transmission/Transmitter/Transmitted

UE: User Equipment

UL: Uplink

UMTS: Universal Mobile Telecommunications System

UTRA: UMTS Terrestrial Radio Access

UTRAN: UTRA Network