Learning-based Coexistence for LTE Operation in Unlicensed Bands

O. Sallent, J. Pérez-Romero, R. Ferrús, R. Agustí
Dept. of Signal Theory and Communications,
Universitat Politècnica de Catalunya (UPC),
Barcelona, Spain
E-mail: [sallent, jorperez, ferrus, ramon]@tsc.upc.edu

Abstract— **The use of Long Term Evolution (LTE) in the unlicensed 5 GHz band, referred to as LTE-U, is a promising enhancement to increase the capacity of LTE networks and meet the requirements foreseen for future systems. Nevertheless, coexistence among several LTE-U and/or Wi-Fi systems in the same band is a key technical challenge to be resolved. In this context, this paper focuses on the channel selection functionality for LTE-U enabled cells to decide the most appropriate channel to use for downlink traffic offloading in the unlicensed band. A distributed Q-learning mechanism that exploits prior experience is proposed to support this functionality, thus enabling coexistence with other systems in a smart and efficient way. The behavior of the proposed approach is illustrated in an indoor scenario with small cells from two different operators. A fully decentralized approach, where the channel selection decision-making is performed independently by each small cell in the scenario, is considered to initially assess the potentials of the Q-learning solution. Promising results are obtained revealing that the throughput achieved by the proposed approach can be between 96% and 99% of the optimum ideal achievable throughput.**

Keywords - LTE-U; Unlicensed bands; Channel Selection, Q-learning; Coexistence.

I. INTRODUCTION

Long Term Evolution Unlicensed (LTE-U) is a promising enhancement in the 3GPP ecosystem that enables LTE to operate and coexist with other technologies in unlicensed bands. Although licensed spectrum remains 3GPP operators' top priority to deliver advanced services and better user experience (i.e., the benefits of licensed spectrum such as controlled Quality of Service (QoS) cannot be matched by unlicensed spectrum), the use of unlicensed spectrum will be an important complement to meet the ultra-high capacity foreseen to be needed by 4G and beyond. Therefore, the view is building LTE-U on the LTE ecosystem in a unified LTE network.

Compared to the usage of Wi-Fi in unlicensed spectrum, LTE-U offers several features that are attractive to operators: (i) The spectrum efficiency and coverage with LTE is better than with Wi-Fi due to more advanced radio features such as robust FEC (Forward Error Correction), hybrid ARQ (Automatic Repeat request), interference coordination/ avoidance, etc., (ii) The same RAN (Radio Access Network) can provide LTE data access in licensed and unlicensed spectrum, (iii) A simplified network management and tracking of KPIs (Key Performance Indicators) through a single RAN can be achieved, (iv) Improved network management and load balancing through tighter integration, (v) 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 (O&M) and QoS issues, (vi) LTE ecosystem kinds of applications (e.g., machine-to-machine, device-to-device, etc.) are exploitable in LTE-U.

In this context, 3GPP has already initiated discussions [1] and work on LTE-U in what is referred to as Licensed Assisted Access (LAA) [2][3]. The introduction and adoption of LTE-U brings a number of challenges to be addressed. As unlicensed spectrum, LTE-U must support fair access of multiple LTE-U and 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.

Although early discussions among players agreed that the core technology should be as much frequency agnostic as possible, a clear focus is placed on unlicensed operation in the 5 GHz band. Despite not existing fully harmonized global regulations, all major markets offer more than 300 MHz available in this band.

Furthermore, given that typical traffic today is asymmetric (e.g. with 8:1 data volume ratio between downlink and uplink [4]), the first focus for LTE-U is on leveraging supplemental downlink capabilities over unlicensed spectrum. In this way, licensed band LTE provides reliable connection for mobility, signaling, voice and data in uplink and downlink, while LTE-U boosts data rates and capacity in downlink. Deployment scenarios can consider both licensed LTE-FDD (e.g., 1.8 GHz) and licensed LTE-TDD (e.g., 2.6 GHz) combined with LTE-U in downlink (e.g., 5 GHz).

The deployment of LTE-U can raise coexistence challenges in both indoor and outdoor scenarios. For example, in a home entertainment environment where a mobile terminal receives streamed video from a small cell via LTE-U and forwards it to a TV using Wi-Fi direct [5], coexistence between both transmissions needs to be ensured. Another example is the deployment of outdoor small cells in a business district that benefit from the aggregation of LTE-U carriers to extend the capacity in the area during certain periods of time.

This work is supported by the Spanish Research Council and FEDER funds under RAMSES grant (ref. TEC2013-41698-R).

In this case, coexistence between multiple operators sharing the same spectrum needs to be ensured.

Under the above framework, this paper focuses on the channel selection functionality to decide the most appropriate channel in the unlicensed band to set-up a LTE-U carrier. In particular, the paper proposes the exploitation of learning mechanisms to support this functionality in a smart way enabling the coexistence with other systems operating in the same band. The applicability of learning to channel selection functionality or, more broadly, spectrum selection, has been investigated in recent years mostly in connection to Cognitive Radio Networks [6]-[9]. In the context of Wi-Fi networks, different channel selection schemes for both uncoordinated and centrally managed scenarios are surveyed in [10], where the possibility of exploiting self-learning experience to support channel selection is identified as an open research issue. While there are some evidences of the benefits that such artificial intelligence-based machine learning mechanisms can bring into the management and operation of wireless networks (see e.g., [11][12]), its applicability to a very practical and realistic use-case such as LTE-U deserves further attention. For example, while [11][12] successfully applied reinforcement learning at a centralized management entity, a multi-operator LTE-U scenario may involve more decentralization in the decision making, which clearly influences the way that a solution is designed.

In this framework, the rest of the paper is organized as follows. Section II discusses the LTE-U coexistence and operation aspects. Then, Section III presents the proposed Q-learning solution for channel selection and Section IV presents some results to illustrate the dynamic behavior of the learning process and to assess the performance of the proposed approach. Finally conclusions are summarized in Section V.

II. LTE-U Coexistence Mechanisms

As stated above, coexistence is a key technical challenge to resolve for the deployment of LTE-U. There exist several mechanisms identified to facilitate coexistence between different systems in the unlicensed 5 GHz. In general terms, these mechanisms exploit the frequency and the time domains, as illustrated in Fig. 1 and detailed in the following subsections.

A. Channel Selection

Channel Selection (also denoted as carrier selection) is the mechanism used to decide the operating channel (i.e. center frequency and associated bandwidth) where a small cell sets up a LTE-U carrier. Therefore, it can be used as a frequency-domain coexistence mechanism to safeguard that LTE is a "good neighbor" in unlicensed bands without requiring modifications in LTE PHY/MAC standards, e.g. just by enabling small cells to choose the cleanest channel based on received power measurements. 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 Rel. 10/11 procedures [13]. This ensures that the interference is avoided between the small cell and its neighboring Wi-Fi devices and/or other LTE-U small cells, provided that there are clean frequencies available. For most Wi-Fi and LTE-U small cell deployments, Channel Selection is usually sufficient to achieve "good neighbor" coexistence [13].

It is worth noting that, 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 of unlicensed devices with radar systems. Specifically, ETSI mandates a Channel Availability Check (CAC) to detect the presence of radar signals in the different channels [14]. Therefore, only channels where no radar signals are detected are declared as available and can be selected for operation by LTE-U or other Wi-Fi systems.

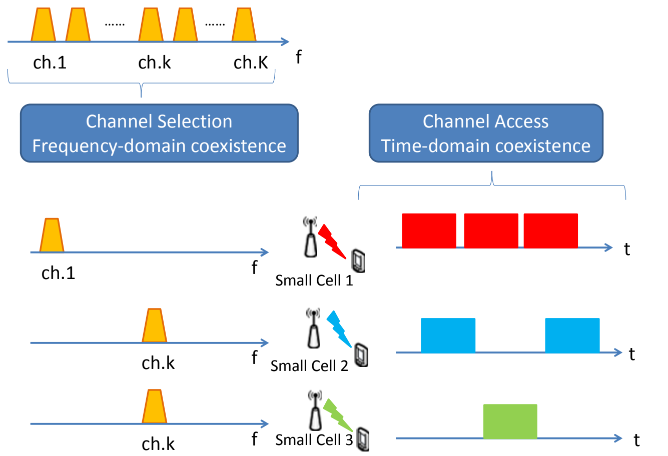

Fig. 1. Illustration of the frequency and time-domain coexistence mechanisms

B. Channel Access

Channel Access is the mechanism used to decide actual transmissions on the selected channel. Channel Access can be used as a time-domain coexistence mechanism to allow that multiple LTE-U small cells and Wi-Fi access points share the same operating channel by carrying out their transmissions in different time instants. This is illustrated in the example shown in Fig. 1 where small cells 2 and 3 have selected the same operating channel *k*. Different strategies may be applied to ensure a fair coexistence in the time domain. In some markets like Europe and Japan regulation requires the support of a Listen-Before-Talk (LBT) scheme that operates at milliseconds scale. For that purpose, following the strategy explained in [14] a small cell using an LTE-U carrier will only transmit if it senses the channel as free during the Clear Channel Assessment (CCA) time (whose duration should be at least 18µs), meaning that the received power in this channel is below a given threshold TL. Then, transmission will be done during a maximum time of 10 ms followed by an idle period θ_{idle} of at least 5% of the channel occupancy time, after which the CCA will be executed again. Notice that, if LBT is required, changes in the LTE air interface will be needed.

In other markets, like US, Korea, India and China, there are no such LBT requirements, but only limitations in the maximum transmit power and out-of-band emissions are specified. In this case, techniques that enable coexistence with Wi-Fi can be actually implemented for LTE Release 10/11 systems without changing the LTE air interface protocol. One example is the Carrier-Sensing Adaptive Transmission (CSAT) [13], which periodically activates and de-activates the LTE-U transmission using LTE MAC control elements to adjust the duty cycle as a function of the measured activity in a channel.

C. LTE-U throughput characterisation

Taking into consideration the abovementioned coexistence mechanisms, this section presents a model to assess the throughput that can be obtained in a LTE-U carrier. For that purpose, let assume a number of small cells denoted as i=1,...,S making use of the 5 GHz unlicensed band as a supplemental downlink for extending the available capacity in the licensed band. The total band is considered to be organized in channels of bandwidth B, numbered as k=1,..,K. Since the focus of the paper is on the coexistence between the different small cells sharing the band, it is assumed that no radar signal is present in any of the K channels.

Considering that the Channel Selection functionality has chosen the k-th channel for carrying out LTE-U transmissions in the downlink of the i-th small cell, and that LBT is required, the total aggregated throughput served by this cell can be estimated as:

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

where $N(i)$ is the total number of users being served by the i-th small cell exploiting the supplemental downlink capacity offered by LTE-U; $SINR_n(i,k)$ 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 (CCA time is already included in these idle periods); and $M(i,k)$ 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 (i.e. those that when they are transmitting they are received above threshold TL at the i-th small cell). As considered in [15], $S(SINR_n(i,k))$ is a generic function ranging between 0 and S_{max} that provides the spectral efficiency in b/s/Hz as a function of $SINR_n(i,k)$ depending on the characteristics of the technology. In turn, $SINR_n(i,k)$ depends on the propagation conditions between the n-th user and the i-th small cell and on the interference generated by other small cells using the k-th channel and that, when they transmit, they are detected at the i-th cell below threshold TL so that they are not sharing the channel in the time domain based on the LBT.

Note that, as a result of LBT, expression (1) assumes an equal sharing in the time domain between small cells, so that on average the i-th small cell can only transmit during a fraction of time $(1-\theta_{idle})/M(i,k)$, where θ_{idle} accounts for the waiting periods imposed by LBT and the fraction $1/M(i,k)$ accounts for the fraction of time that the i-th small cell can transmit because it senses the k-th channel as free. It is also assumed that all the small cells operate only with LTE-U in the downlink direction. Similarly, (1) assumes a full buffer traffic model in which the small cell always has data to be transmitted, and that the total bandwidth B is equally shared between all the $N(i)$ users being served by the i-th small cell, so that on average a user observes a fraction $B/N(i)$ of the total bandwidth. It is worth mentioning that expression (1) could be easily modified to capture other scheduling strategies to share the bandwidth between users. Note also that (1) corresponds to the throughput achievable in one channel. In case that a small cell aggregates multiple channels, the total throughput would be the summation of (1) for all the channels.

III. LEARNING-BASED CHANNEL SELECTION IN LTE-U

Clearly, the design of a proper Channel Selection functionality can greatly improve the overall efficiency of the LTE-U operation. Specifically, it can be derived from (1) that the decision-making applied to perform the channel selection for the i-th small cell will impact on the achieved throughput performance mainly through the terms $M(i,k)$ and $SINR_n(k)$. Thus, if the selected k-th channel is not used by other cells (i.e., $M(i,k)$=1), higher throughput will follow. Similarly, if the selected k-th channel is affected by low interference levels, high $SINR_n(k)$ will be observed and higher throughputs will follow. Therefore, the channel selection for a given small cell should be able to dynamically identify and capture the relevant context information about the current status of utilization of the candidate channels so that the most adequate ones can be selected. Consequently, smart solutions able to identify the best channels under each specific condition are of high interest for the materialization of all the potentials that LTE-U offers.

From an architectural point of view, different approaches for Channel Selection can be envisaged: (a) fully distributed case, where each small cell makes decisions on its own, (b) intra-operator coordination, where decisions for a given small cell take into consideration knowledge about other small cells' configurations belonging to the same operator, (c) inter-operator coordination, where also information about small cells from other operators in the area is available and (d) coordination also with managed Wi-Fis in the area. Notice that unmanaged legacy Wi-Fi unable to explicitly provide information about its configuration may also be present in the scenario. Clearly, higher coordination levels will ease the Channel Selection decision-making. However, higher coordination levels involve more demanding network coordination architectures, information exchange protocols and procedures, etc. In this respect, while the purpose of this work is not to deepen into suitable architectures and comparative trade-off analysis, this paper shifts the focus towards a fully decentralized approach and explores to what extent the inclusion of a smart Channel Selection logic can overcome the intrinsic disadvantages associated with the fact that no explicit knowledge about the other small cells and/or Wi-Fis operating in the area is available.

From a decision-making logic point of view, exploiting learning from past experience seems a pertinent principle in LTE-U context. Indeed, given that the initially envisaged LTE-U scenarios (either indoor or outdoor) are infrastructure-based and, consequently, rather stable (in the sense that there will be a number of small cells deployed around that remain operative for months/years together with possible additions/relocations of small cells at a long-term time scale), each small cell may autonomously learn what channels are usually not being used by its neighbors and then tend to select such free channels. Thanks to this learning capability, general scanning procedures over the 5 GHz band conducted systematically to look for the cleanest channel can be avoided or reduced to a minimum. Besides, learning from the own experience in using a channel can help in overcoming situations like the hidden node problem where a small cell can detect a channel as not used but some of its served terminals can experience severe interference conditions from other small cells and/or Wi-Fis. Furthermore, 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 (e.g., the deployment of a new small cell in the area).

With all the above, this paper proposes the use of a Q-learning solution as an efficient means to carry out a distributed Channel Selection in a practical while at the same time efficient way. Q-learning belongs to the category of Temporal Difference Reinforcement Learning (RL) techniques that consist in learning how to map situations to actions so as to maximize a scalar reward [16]. The learning is achieved through the interaction with the environment, so that the learner discovers which actions yield the most reward by trying them. In this way, the idea proposed by this paper is that each small cell progressively learns and selects the channels that provide the best performance based on the previous experience.

In particular, in the proposed approach 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 a single state Q-learning approach with null discount rate given by [16]:

$$Q(i,k) \leftarrow (1-\alpha_L)Q(i,k)+\alpha_L \cdot r(i,k) \tag{2}$$

where $\alpha_L \in (0,1)$ 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 maximizes the total throughput, the reward function considered in this paper is given by:

$$r(i,k)=\frac{\overline{R(i,k)}}{R_{\max}} \tag{3}$$

where $\overline{R(i,k)}$ 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. In turn, $R_{\max}=B \cdot S_{\max} \cdot (1-\theta_{idle})$ is a normalization factor.

At initialization, i.e. when channel k has never been used in the past by small cell i, $Q(i,k)$ is set to an arbitrary value Q_{ini}.

Based on the $Q(i,k)$ value functions, the proposed Channel Selection decision-making for the small cell i follows the softmax policy [16] in which channel k is chosen with probability:

$$\Pr(i,k)=\frac{e^{\frac{Q(i,k)}{\tau(i)}}}{\sum_{k'=1}^{K} e^{\frac{Q(i,k')}{\tau(i)}}} \tag{4}$$

where $\tau(i)$ is a positive parameter called *temperature*. High temperature values cause 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 making is a popular means of balancing the exploitation and exploration dilemma in RL-based schemes. It exploits what the system already knows in order to obtain reward (i.e. selecting with high probability those channels that have provided good results in the past), but it also explores in order to make better actions in the future (i.e. the selection must try first a variety of channels and progressively favor those that appear to be the best ones) [16]. A cooling function will be considered in this paper to reduce the value of the temperature $\tau(i)$ 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. Specifically, the following logarithmic cooling function is assumed [17]:

$$\tau(i)=\frac{\tau_0}{\log_2(1+n(i))} \tag{5}$$

where 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.

IV. Performance Analysis

This section presents some evaluation results to illustrate the behavior and the performance of the proposed approach.

A. Simulation scenario

The considered scenario is based on the indoor scenario for LTE-U coexistence evaluations defined in the context of the corresponding 3GPP Study Item [3]. It consists of a single floor building where two operators deploy 4 small cells (SCs) each. SCs are equally spaced and centered along the shorter dimension of the building, as depicted in Fig. 2. Small cells SC1 to SC4 are owned by operator 1 (OP1), while SC5 to SC8 are owned by operator 2 (OP2). SCs are deployed at height 6m while the antenna height of the mobile terminals is 1.5m. A total of 10 terminals (users) per operator are randomly distributed inside the building. Each user is associated to the SC of its own operator that provides the highest received power. The SC-to-terminal and SC-to-SC path loss and shadowing are computed using the ITU InH model in [18]. 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 function $S(SINR)$ is obtained from Section A.1 in [15] with S_{max}=4.4 b/s/Hz.

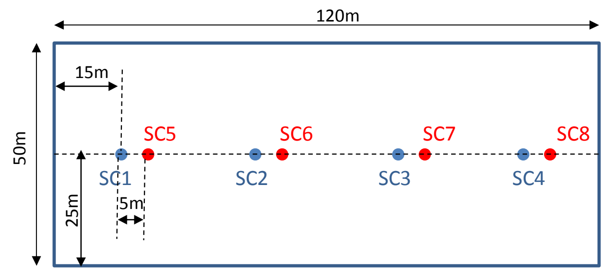

Fig. 2. Layout of the floor building

The threshold TL used in the CCA of the LBT to decide whether a channel is sensed as free or not is TL=-70 dBm/MHz according to the formula in [14]. With this threshold and the considered propagation model it turns out that in the layout shown in Fig. 2 only SC3 and SC6 are able to detect the transmissions of all the other small cells. On the contrary, SC1 is not able to detect the transmissions of SCs 4, 7 and 8; SC2 does not detect SC8; SC4 does not detect SCs 1 and 5; SC5 does not detect SCs 4 and 8; SC7 does not detect SC1; and SC8 does not detect SCs 1, 2 and 5. Besides, θ_{idle}=0.05 is assumed. The parameters related to the Q-learning algorithm are α_L=0.1, τ_0=0.15 and Q_{ini}=0.5.

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

B. *Analysis of the Q-learning behavior*

To gain insight into the behavior of the proposed Q-learning approach for channel selection, different situations are analyzed in this sub-section, depending on the number of available channels and on the selection strategy followed by each operator.

Firstly, let consider the case in which there are K=8 available channels. OP2 is assumed to have a fixed frequency assignment in which SC5 to SC8 make use of channels k=5 to k=8, respectively. OP1 has no knowledge about OP2 deployment and selected channels, and it applies the Q-learning approach in a decentralized way for SC1 to SC4. Fig. 3 depicts an example of evolution of the channel selection probabilities for the SCs of OP1. As it can be observed, after an initial phase in which different channels are explored, each SC learns to select one channel with probability close to 1. The solution learnt in this example corresponds to using the channels k=1,2,4,3 respectively for SC1, SC2, SC3 and SC4. Clearly, this corresponds to one of the optimal solutions since each channel is used only by one of the SCs in the scenario, time domain sharing is avoided and therefore throughput is maximized. It can be also observed in Fig. 3 that the duration of the initial exploratory phase before converging to the learnt solution is in the order of 1500 time steps. Then, considering that each small cell generates activity periods every 150 time steps on average, this means that convergence is achieved after performing about 10 channel selection decisions.

Secondly, in order to illustrate the capability of Q-learning to adapt to changes in the scenario, let consider a variation of the previous situation. In this case, at the beginning of the simulation SC7 is switched off, and then it is switched on using channel k=7 after t=30000 time steps. Results in terms of channel selection probabilities for the SCs of OP1 are depicted in the example shown in Fig. 4. It can be observed that, while SC7 of OP2 is switched off and channel k=7 is not used by OP2, the solution learnt by SCs 1 to 4 corresponds to using channels k=2,3,7,1, respectively, again avoiding sharing the same channel in the time domain. However, when SC7 switches on at t=30000 and starts using channel k=7, SC3 learns that this channel is no longer a good solution, so the probability of selecting this channel starts to decrease. Eventually, SC3 learns to use channel k=4, so that again each small cell is using a different channel.

Thirdly, the situation when the SCs of both operators apply the Q-learning strategy to select among K=8 channels is analyzed. Fig. 5 depicts the evolution of the selection probabilities. It can be observed that, after some initial exploratory phase, the solution found is that SC1 to SC4 from OP1 learn to select channels k=4,6,2,1, respectively, while SC5 to SC8 from OP2 learn to select channels k=5,3,7,8, which also corresponds to an optimum solution in which there is no channel reuse. It is also worth noting that the exploratory phase for some SCs in this case takes longer than in the previous cases. This is because 8 SCs are intending to learn from and act on the environment in this case, compared to the previous case where only 4 SCs were applying the Q-learning strategy.

Let now analyze the more challenging case of K=4 available channels. Therefore, sharing the same channels in the time domain is unavoidable. Fig. 6 depicts an example with the evolution of the selection probabilities for the case when SC1 to SC4 apply Q-learning while OP2 uses a fixed assignment of the four available channels k=1,2,3,4 in the small cells SC5 to SC8, respectively. In addition, SC7 using k=3 is switched off at the beginning of the simulation and it switches on at t=30000

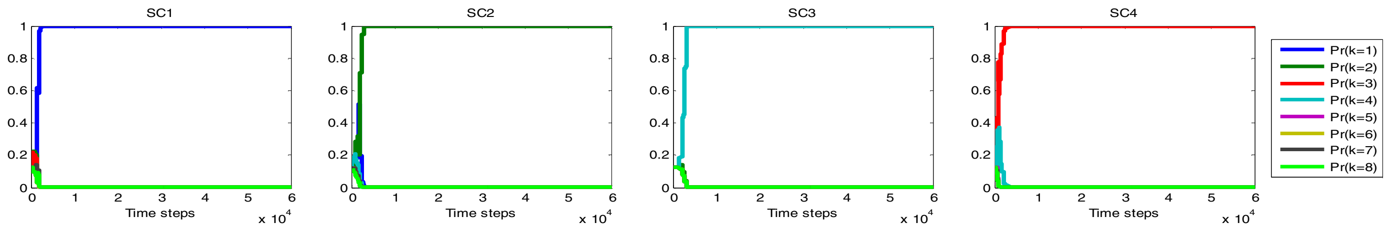

Fig. 3. Evolution of the channel selection probabilities for the SCs of operator 1 with K=8 channels and Operator 2 using a fixed assignment.

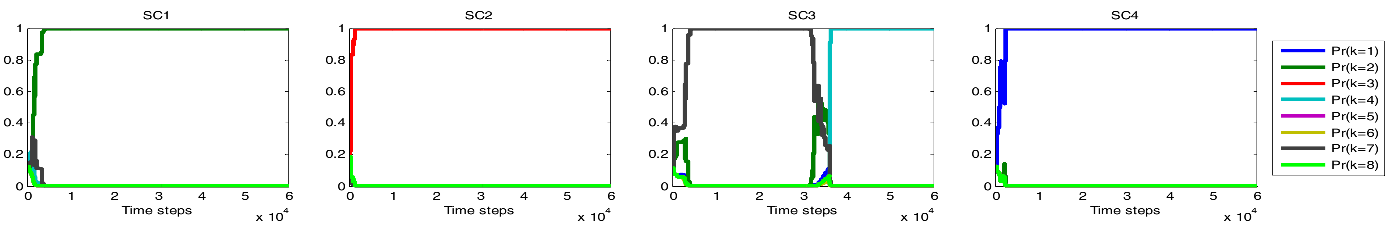

Fig. 4. Evolution of the channel selection probabilities for SCs of operator 1 when a new SC is activated at t=30000.

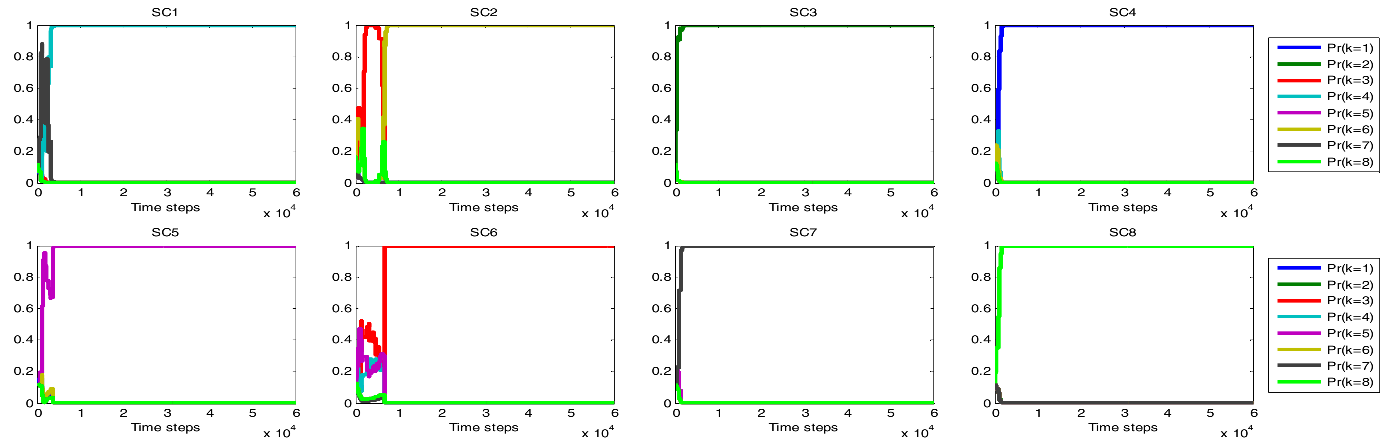

Fig. 5. Evolution of the channel selection probabilities with K=8 channels when both operators apply Q-learning.

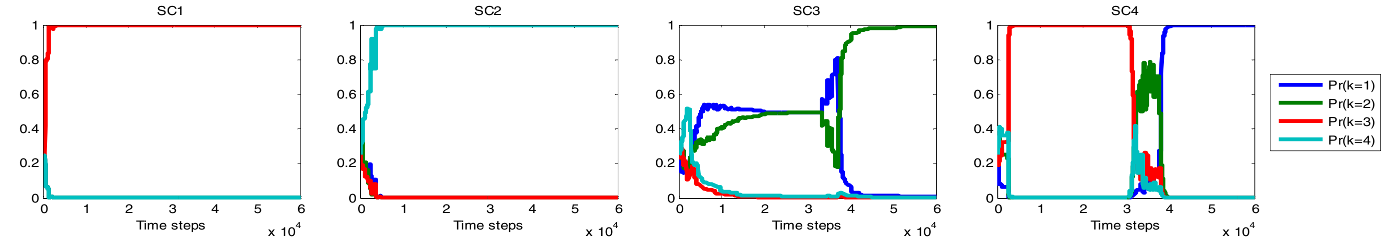

Fig. 6. Evolution of the channel selection probabilities for the SCs of operator 1with K=4 channels when a new SC is activated at t=30000.

time steps. It can be observed how the initial solution learnt until t=30000 reflects that SC1 and SC4 learn to use channel k=3 (i.e. the only channel not used by OP2 at the beginning). This is certainly a good choice, since SC1 and SC4 are located far away from one another inside the building (see the layout in Fig. 2) so they are not mutually detected during the CCA phase (i.e., the received power level is below the TL threshold). Therefore, they can use the same channel without sharing it in the time domain. Similarly, SC2 learns to use the same channel k=4 as SC8, which is also located at a sufficiently large distance from SC2. As for SC3, it is located in the middle of the building and it is able to detect all the other SCs during the CCA phase, so it necessarily has to share a channel in the time domain following LBT. As seen in Fig. 6, it learns not to use channels k=3 and 4, which are already in use by two other SCs. Instead, it can use indifferently either channel k=1 or 2 as both of them are used only by another SC. As both channels provide the same performance, they are selected with equal probability.

After t=30000, SC7 is switched on and starts using channel k=3, which is the channel that was being used by SC1 and SC4 so far. Since SC1 is located at a sufficient distance from SC7 in order not to detect it during CCA, it keeps on operating in the same channel k=3, as seen in Fig. 6. Instead, SC4, located close to SC7, perceives throughput degradation in this channel due to the time sharing between SC4 and SC7, and progressively reduces its selection probability. At the end, SC4 learns to use the same channel k=1 as SC5 that is located at the other side of the building. With all these changes, also SC3 identifies that channel k=1 is no longer a good option (since it is used by both SC4 and SC5) and it learns to use channel k=2. The final solution learnt by the small cells of OP1 is the use of channels k=3,4,2,1, respectively by SC1 to SC4. A detailed analysis of the throughput achievable with all the possible solutions (not reported for the sake of brevity) reveals that this is actually the solution that maximizes the total aggregated throughput among all the SCs of both operators.

C. *Performance analysis*

In order to assess the benefits of the proposed solution from a quantitative perspective, Fig. 7 plots the ratio between the total throughput achieved by the proposed Q-learning in the scenario with respect to the throughput that would be achieved with an ideal optimum assignment. This optimum solution corresponds to the assignment of channels to small cells that maximizes the total throughput in the scenario. Results are provided for the cases K=4 and K=8 channels, and for the cases when Q-learning is applied only by OP 1 (OP 2 following a fixed assignment) or by the two operators. Each result corresponds to the average throughput obtained after 50 different experiments each one associated with a different user spatial distribution. Each experiment lasts for 1E6 time steps and the throughput for all SCs is aggregated and averaged along the whole simulation time. In each experiment the optimum channel assignment is found by exhaustive analysis of all the possible combinations.

As it can be observed in Fig. 7, the proposed approach achieves between 95.8 and 98.8% of the optimum throughput, revealing a promising behavior of the Q-learning mechanism. Reasonably, slightly better results are achieved with K=8 than with K=4, which is a more challenging scenario. Similarly, slightly better results are achieved when both operators apply Q-learning than when only one is doing so. In turn, Fig. 8 depicts the Cumulative Distribution Function (CDF) of the throughput (normalized to R_{max}) obtained for all the considered experiments. The figure considers that both operators apply Q-learning with K=4 channels. It can be noticed how the performance of the Q-learning approach is very close to the optimum one. As a further reference, the fully random case where each SC selects randomly the channel to be used is also presented, revealing that the Q-learning approach offers a very significant improvement with respect to this basic strategy.

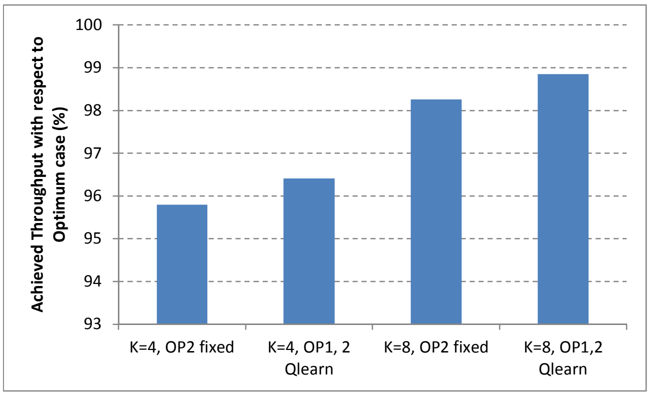

Fig. 7. Throughput achieved by Q-learning with respect to the optimum case.

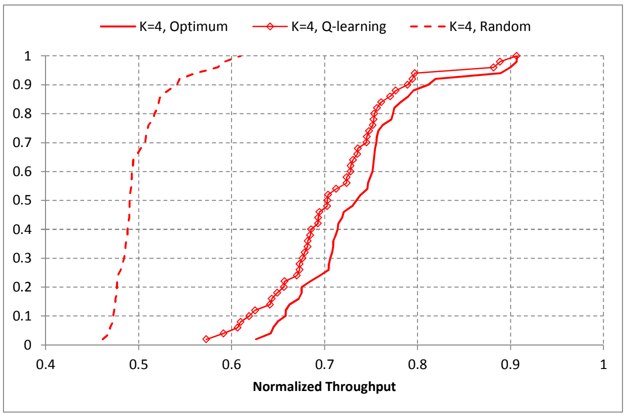

Fig. 8. CDF of the achieved normalized throughput for K=4 when the two operators apply Q-learning.

V. Conclusions and Future Work

The use of LTE-U in the unlicensed 5 GHz band is a promising enhancement to meet the requirements foreseen for future systems. Coexistence between different systems operating in the same band is one of the key technical challenges to be resolved for a successful operation of LTE-U deployments. In this framework, this paper has addressed the Channel Selection functionality that decides the most appropriate channel in the unlicensed band to set-up a LTE-U carrier for supplemental downlink as a means to facilitate the coexistence. In particular, a distributed Q-learning mechanism that exploits prior experience has been proposed. In order to initially assess the potentials of the Q-learning solution, a fully decentralized approach has been considered. The evaluations presented in an indoor scenario with small cells belonging to different operators have revealed promising results in which the proposed approach is able to achieve a performance between 96% and 99% of the optimum ideal achievable throughput.

Based on these promising results, a number of areas are identified for further consolidating the proposed approach. In particular, different intra and inter-operator coordination levels can be studied both in terms of architectural implications and Q-learning Channel Selection strategy design. Similarly, the study should be extended to include a rigorous stability analysis of the Q-learning approach, particularly when legacy unmanaged Wi-Fi with unknown channel selection strategies are in place, and when different activity conditions exist in the different small cells. Finally, the combination of the proposed approach with other possible solution approaches such as Game Theory can also be developed.

References

[1] 3GPP workshop on LTE in unlicensed spectrum, Sophia Antipolis, France, June 13, 2014. http://www.3gpp.org/ftp/workshop/2014-06-13_LTE-U/

[2] RP-141664, Ericsson, Qualcomm, Huawei, Alcatel-Lucent, “Study on Licensed-Assisted Access using LTE”, 3GPP TSG RAN Meeting #65, Edinburgh (Scotland), 9 - 12 September 2014.

[3] 3GPP TR 36.889, “Study on Licensed-Assisted Access to Unlicensed Spectrum (Release 13)”, November, 2014.

[4] S. Nielsen, A. Toskala, “LTE in Unlicensed Spectrum: European Regulation and Co-existence Considerations”, 3GPP workshop on LTE in unlicensed spectrum, Sophia Antipolis, France, June 13, 2014.

[5] RWS-140010, Sony, “Requirements and Proposed Coexistence Topics for the LTE-U Study”, 3GPP workshop on LTE in unlicensed spectrum, Sophia Antipolis, France, June 13, 2014.

[6] I. Macaluso, D. Finn, B. Ozgul, L. A. DaSilva, “Complexity of Spectrum Activity and Benefits of Reinforcement Learning in Dynamic Channel Selection”, IEEE Journal on Selected Areas in Communications, Vol. 31, No. 11, November, 2013.

[7] J. Pérez-Romero, O. Sallent, R. Agustí, “Enhancing Cellular Coverage through Opportunistic Networks with Learning Mechanisms”, GLOBECOM, December, 2013

[8] S. Chen, R. Vuyyuru, O. Altintas, A.M. Wyglinski, “Learning-based channel selection of VDSA Networks in Shared TV Whitespace”, VTC Fall Conference, Quebec, Canada, September, 2012.

[9] Y. Li, H. Ji, X. Li, V.C.M. Leung, “Dynamic channel selection with reinforcement learning in cognitive WLAN over fiber”, International Journal of Communication Systems, March, 2012.

[10] S. Chieochan, E. Hossain, J. Diamond, “Channel Assignment Schemes for Infrastructure-Based 802.11 WLANs: A Survey”, IEEE Communications Surveys and Tutorials, Vol. 12, No. 1, 2010.

[11] F. Bernardo, R. Agustí, J. Pérez-Romero, O. Sallent, “An Application of Reinforcement Learning for Efficient Spectrum Usage in Next Generation Mobile Cellular Networks”, IEEE Transactions on Systems, Man and Cybernetics - Part C, Vol. 40, No. 4, pp. 477-484, July, 2010.

[12] N. Vucevic, J. Pérez-Romero, O. Sallent, R. Agustí “Reinforcement Learning for Joint Radio Resource Management in LTE-UMTS Scenarios”, Computer Networks, Elsevier, May, 2011, Vol. 55, No. 7, pp. 1487-1497.

[13] Qualcomm, “LTE in Unlicensed Spectrum: Harmonious Coexistence with Wi-Fi”, June 2014.

[14] ETSI EN 301 893 v1.7.2 “Broadband Radio Access Networks (BRAN): 5 GHz high performance RLAN; Harmonized EN covering the essential requirements of article 3.2 of the R&TTE Directive”, July, 2014.

[15] 3GPP TR 36.942 v12.0.0, “Radio Frequency (RF) system scenarios”, September, 2014.

[16] R.S. Sutton, A. G. Barto, Reinforcement Learning: An Introduction, MIT Press, 1998

[17] S. Geman, D. Geman, “Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images”, IEEE Transactions on Patterns Analysis and Machine Intelligence, Vol. PAMI-6, No. 6, November, 1984, pp. 721-741.

[18] 3GPP TR 36.814 v9.0.0 “Further advancements for E-UTRA physical layer aspects”, March, 2010.