IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received February 2, 2018, accepted April 16, 2018, date of publication May 15, 2018, date of current version June 5, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2829492

A Q-Learning Scheme for Fair Coexistence Between LTE and Wi-Fi in Unlicensed Spectrum

VASILIS MAGLOGIANNIS, DRIES NAUDTS, ADNAN SHAHID, (Senior Member, IEEE), AND INGRID MOERMAN
IDLab, Department of Information Technology, Ghent University–imec, B-9052 Ghent, Belgium

Corresponding author: Vasilis Maglogiannis (vasilis.maglogiannis@ugent.be)

This work was supported by the European H2020 Programs through the eWINE Project under Grant 688116 and through the WiSHFUL Project under Grant 645274.

ABSTRACT During the last years, the growth of wireless traffic pushed the wireless community to search for solutions that can assist in a more efficient management of the spectrum. Toward this direction, the operation of long term evolution (LTE) in unlicensed spectrum (LTE-U) has been proposed. Targeting a global solution that respects the regional regulations worldwide, 3GPP has published the LTE licensed assisted access (LAA) standard. According to LTE LAA, a listen before talk (LBT) procedure must precede any LTE transmission burst in the unlicensed spectrum. However, the proposed standard may cause coexistence issues between LTE and Wi-Fi, especially in the case that the latter does not use frame aggregation. Toward the provision of a balanced channel access, we have proposed mLTE-U that is an adaptive LTE LBT scheme. According to mLTE-U, LTE uses a variable transmission opportunity (TXOP), followed by a variable muting period. This muting period can be exploited by co-located Wi-Fi networks to gain access to the medium. In this paper, the system model of the mLTE-U scheme in coexistence with Wi-Fi is studied. In addition, mLTE-U is enhanced with a Q-learning technique that is used for autonomous selection of the appropriate combinations of TXOP and muting period that can provide fair coexistence between co-located mLTE-U and Wi-Fi networks. Simulation results showcase the performance of the proposed model and reveal the benefit of using Q-learning for self-adaptation of mLTE-U to the changes of the dynamic wireless environment, toward fair coexistence with Wi-Fi. Finally, the Q-learning mechanism is compared with conventional selection schemes showing the superior performance of the proposed model over less complex mechanisms.

INDEX TERMS LTE, Wi-Fi, Q-learning, fairness, coexistence, LTE unlicensed, transmission opportunity, muting period.

I. INTRODUCTION

Over the last years, the technological growth has led to a tremendous increase of wireless devices such as smartphones, laptops and sensor networks, that exchange information with each other. Additionally, the establishment of Internet of Things (IoT) has further increased the number of the wirelessly interconnected devices. The wireless traffic is expected to increase by a factor of 1000 by 2020 compared to that in 2010 [1]. This information is exchanged between devices using different types of wireless technologies such as LTE, IEEE 802.11 (also known as Wi-Fi), IEEE 802.15.4 and Bluetooth. Recently, technologies that target wide range communications such as LORA and SIGFOX exploit sub-GHz bands. Furthermore, high frequency bands such as mmWave are used for mutli-gigabit speeds (IEEE 802.11ad). It is clear that soon the wireless network capacity will become a bottleneck for serving the increased wireless traffic.

Concurrently, the licensed spectrum used by the mobile operators becomes very scarce. The availability of the licensed spectrum combined with the high cost of a licensed frequency band have pushed the mobile operators to investigate solutions that can assist in meeting the 1000x challenge requirements. Among other solutions like (massive) Multiple-Input Multiple-Output (MIMO) and Carrier Aggregation the LTE operation in the unlicensed spectrum (LTE-U) has attracted significant attention from the wireless community. Hence, several techniques have been proposed aiming to achieve harmonious coexistence between LTE and other well-established technologies in the unlicensed spectrum (e.g. Wi-Fi) [2].

In regions where a Listen Before Talk (LBT) procedure before a transmission is not mandatory, such as in U.S.A. or in China, LTE can transmit in unlicensed spectrum using a duty-cycle technique. The most famous technique of this nature is the Carrier Sense Adaptive Transmission (CSAT) [3], which has been proposed by Qualcomm. CSAT exploits duty-cycle periods in order to give transmission opportunities (TXOP) to potential co-located Wi-Fi networks. According to CSAT, the time domain is divided into ON and OFF periods. During an OFF period, also known as mute period, LTE remains silent, giving the opportunity to other networks to transmit. During an ON period, LTE accesses the channel without estimating it for potential ongoing transmissions. The duration of the LTE ON and OFF periods are defined by the evolved NodeB (eNB) according to the observed channel utilization, based on the estimated number of Wi-Fi Access Points (AP) [3].

Towards a coexistence technique that respects the regional regulations in regions where an LBT procedure before a transmission in the unlicensed spectrum is mandatory (such as in Europe and in Japan), 3GPP published the LTE License Assisted Access (LTE LAA) standard as part of the Release 13 [4]. The standard includes the description of an LBT procedure that is also known as Clear Channel Assessment (CCA) that must be performed prior to a transmission in the unlicensed spectrum. Initially, LTE LAA is designed to be used for downlink (DL) traffic only and to operate within the 5-GHz unlicensed band. In a latter phase and towards Release 14, it is expected to be used for both DL and uplink (UL) traffic [5]. According to LTE LAA, an eNB will be able to activate and deactivate a secondary cell in unlicensed spectrum, next to the primary cell that operates in the licensed band owned by the operator. Through a secondary cell, an operator can offload the LTE network by transmitting DL data traffic via the Physical DL Shared Chanel (PDSCH), while the LTE control signals and the UL traffic (according to Release 13) will be transmitted via the licensed anchor. Furthermore, the LTE operation solely in the unlicensed spectrum has been proposed by leading wireless stakeholders, towards the decoupling of LTE from the operators. To this end, they formed the MulteFire Alliance [6].

Although the LTE LAA standard defines that a CCA procedure must be performed before a transmission in the unlicensed spectrum, it also defines four channel access priority classes. Each priority class specifies among others the transmission duration in unlicensed channel after it has been estimated as idle. This transmission duration varies from 2 up to 10 ms. On the other hand, when frame aggregation is not enabled or supported by the 802.11 standard, a typical Wi-Fi transmission lasts for few hundreds of μs [7]. Even when frame aggregation is used, a significant percentage of packets requires a short transmission time. In [8], it has been evaluated that 50% of the packets are transmitted within 30 μs, while 80% of the packets are transmitted within 1 ms. This shows that the ratio between LTE and Wi-Fi transmission time occupancy is not balanced. This can lead to unfair coexistence between the two networks in the unlicensed spectrum.

In our previous work [9] and based on this observation, a novel coexistence mechanism named mLTE-U has been proposed and builds on elements of LTE Release 13. mLTE-U is an adaptive LTE-U transmission scheme, according to which LTE can transmit DL traffic in the unlicensed spectrum after the channel has been assessed as idle, using a variable TXOP period followed by a variable muting period. This muting period can give channel access opportunities to other potentially co-located networks such as Wi-Fi. From the different possible pairs of TXOPs and muting periods, the selection of the appropriate combination has to be done in a way that the co-located networks share the medium in a fair way. The mLTE-U scheme has been evaluated using an event-based simulation platform.

This article further extends this work by studying the system model of the mLTE-U mechanism in coexistence with Wi-Fi and by introducing reinforcement learning and specifically Q-learning, as it is able to provide automatic and autonomous selection of the appropriate TXOP and muting period combinations that can enable fair coexistence. Q-learning is a technique that converges to optimal policies. Another advantage of Q-learning is that it does not require a prior environment model [10]. This is suitable for dynamic and arbitrary environments such as wireless environments. The main contribution of this work is summarized as follows:

- Description and analysis of the system model for the proposed mLTE-U scheme when it coexists with Wi-Fi or other mLTE-U networks
- Discussion about fair coexistence in unlicensed spectrum, definition of fairness as equal sharing of spectrum in a technology-agnostic way and problem formulation of mLTE-U TXOP and muting period selection towards fair spectrum sharing
- Use of Q-learning mechanism for optimal and autonomous selection of mLTE-U TXOP and muting period towards fair coexistence
- Performance evaluation of the proposed mLTE-U coexistence scheme with and without using Q-learning mechanism through simulations
- Comparison of Q-learning with conventional selection mechanisms such as random selection and round-robin

The remainder of the article is organized as follows. Section II gives an overview of the current literature on the coexistence of LTE-U and Wi-Fi and the exploitation of Q-learning towards the selection and adjustment of coexistence parameters. In Section III, we discuss the problem that arises when LTE LAA coexists with traditional Wi-Fi networks that do not use frame aggregation and we give a summarized description of the mLTE-U scheme. Next, in Section IV, we analyze the system model of the mLTE-U scheme, when it coexists with Wi-Fi. Section V discusses the topic of fair coexistence in unlicensed spectrum and the approach followed in this article. Section VI analyses the integration and usage of a Q-learning mechanism in mLTE-U towards autonomous and optimal selection of the mLTE-U parameters. In Section VII, we describe the simulation environment that has been used, while Section VIII evaluates the performance of the proposed technique and compares it with conventional selection schemes. Finally, Section IX concludes the paper and discusses plans for future work.

II. RELATED WORK

A. COEXISTENCE BETWEEN LTE-U AND Wi-Fi

From the moment LTE-U was firstly introduced, there were serious concerns from the wireless community about unfair coexistence of LTE with other well-established technologies in the unlicensed spectrum, such as Wi-Fi. These concerns were based on the fact that LTE is designed to be a scheduled technology that does not use a CCA mechanism to sense the medium before a transmission. Hence, it would transmit arbitrarily forcing the other networks to continuously back-off. In our previous work [11], we studied the impact of a traditional LTE operating in unlicensed spectrum on Wi-Fi using Off-The-Shelf (OTS) hardware equipment at the LTE testbed of IMEC [12]. Three different levels of LTE signal power have been examined that represent different possible levels of LTE impact on Wi-Fi. According to the results, the Wi-Fi performance can be significantly affected by LTE. Several other studies [13] [14] [15] evaluate the impact of LTE on Wi-Fi through experiments, mathematical models and simulations, all coming to the same conclusion, namely that coexistence mechanisms are required to render LTE fair towards other co-located technologies, like Wi-Fi.

Lately, several coexistence mechanisms have been proposed, targeting to improve the coexistence between LTE and Wi-Fi. Similar to the CSAT mechanism that is described in Section I, Almeida *et al.* [16] propose a coexistence scheme that exploits periodically blank LTE subframes during an LTE frame in order to give transmission opportunities to Wi-Fi. The scheme is evaluated via simulations and it is concluded that the number and the order of the blank subframes have an impact on the provided coexistence.

In our previous work [17], the concept of LTE-U has been extensively studied. To this end, a detailed analysis of the current state-of-the-art regarding LTE-U and Wi-Fi is given. Additionally, a classification of techniques that can be applied between co-located LTE and Wi-Fi networks is presented. This classification combined with the study of the literature revealed the lack of cooperation schemes among co-located networks that can lead to more optimal use of the available spectrum. In order to fill this gap, we proposed several concepts of cooperation techniques that can enhance the spectral efficiency between coexisting LTE and Wi-Fi networks. The proposed techniques are compared between each other in terms of complexity and performance.

As it has been discussed in Section I, 3GPP announced the LTE LAA as part of Release 13, towards a global coexistence technique that respects the regional regulations worldwide. The strong point of this technique is that it includes the description of a CCA procedure that must be performed before a transmission in the unlicensed spectrum to verify the availability of the channel [4]. The concept of the adoption of a CCA procedure by LTE has been proposed in several works. Kim *et al.* [18] propose an LBT scheme for LTE LAA that enhances the coexistence with Wi-Fi and increases the overall system performance. The scheme comprises of two parts named on-off adaptation for channel occupancy time and short-long adaptation for idle time. According to the first mechanism, the channel occupancy time of LTE can be adapted based on the load of the network, while according to the second one the idle period can be adapted based on the Contention Window (CW) duration of Wi-Fi.

Bhorkar *et al.* [19], propose a MAC layer for LTE-U that uses LBT and channel reservation packets. The LBT can be either synchronous or asynchronous. Furthermore, in order to cope with potential collisions, they propose improvements to the LTE link adaptation algorithm. The simulation results show that the performance of co-located Wi-Fi can be improved by the proposed MAC design. The LTE-U cell edge performance can be also improved by the channel reservation mechanism.

Hao *et al.* [20] study the coexistence between LTE LAA and Wi-Fi using LBT Category 4 (Cat 4) channel access scheme. The behavior of the eNB is modeled as a Markov chain. The authors adopt the obtained throughput as performance indicator. The proposed LBT scheme uses an adaptive CW size for LTE LAA. The results show that the proposed scheme can achieve higher performance compared to the fixed CW size scheme.

Mushunuri *et al.* [21] propose an LBT mechanism for LTE LAA that aims to share the medium in a fair way and concurrently to increase of the overall system performance. This work analyses mathematically the proposed LBT scheme and additionally, it is validated via simulations. The results show that the performance of Wi-Fi can be increased by proper selection of LAA channel occupancy and the backoff counter.

A detailed survey of the coexistence between LTE and Wi-Fi on 5 GHz with corresponding deployment scenarios is given in [22]. The authors give a detailed description of the coexistence-related features of LTE and Wi-Fi, the challenges, the differences in performance between the two different technologies and co-channel interference. They discuss in detail the proposed coexistence techniques between LTE and Wi-Fi that have been proposed in the literature. Moreover, the survey analyses the concept of scenario-oriented coexistence. According to this concept, coexistence-related problems can be solved according to different deployment scenarios.

B. COEXISTENCE ENHANCEMENT WITH Q-LEARNING

Q-learning has been used in various works to enhance the coexistence mechanisms and render them capable to learn individually the best possible strategies in order to achieve a target Li *et al.* [23] propose a Q-learning-based dynamic duty cycle selection mechanism for the configuration of LTE transmission gaps. LTE LAA and Wi-Fi performance using a fixed transmission gap is evaluated and is used as reference scenario. Then, the proposed Q-Learning mechanism is compared with the reference scenario. Simulation results

TABLE 1. Channel access priority class configuration of LTE LAA.

Channel access priority class (p)	m_p	$CW_{min,p}$	$CW_{max,p}$	$T_{m\ cot,p}$	Allowed CW_p sizes
1	1	3	7	2 ms	3,7
2	1	7	15	3 ms	7,15
3	3	15	63	8 or 10 ms	15,31,63
4	7	15	1023	8 or 10 ms	15,31,63,127,255,511,1023

show that the proposed scheme enhances the overall capacity performance.

Rupasinghe and Guvenc [24] propose a fair DL traffic management scheme. This scheme targets to adapt the minimum CW values and assign feasible weights to the LAA eNBs with different traffic loads. This way, they aim to achieve fair spectrum sharing with coexisting Wi-Fi networks and service differentiation for DL LTE LAA traffic. Simulation results show that the proposed scheme can offer fair coexistence with Wi-Fi networks and can provide proportional fairness to LAA eNBs with different traffic requirements.

In [25], a docitive Q-learning scheme for joint resource allocation and power control is proposed. In this scheme, the femto base stations learn the optimal strategies by exploiting Q-learning and share their knowledge with their neighbors. The target of the learning scheme is the maximization of the femtocell capacity, while maintaining the quality of service requirement of the macro-users. The proposed scheme is compared with the independent learning in terms of convergence, min-max capacity and the impact on the femtocell density.

A channel selection mechanism using Q-learning for LTE-U is proposed in [26]. This mechanism decides the most appropriate channel in unlicensed spectrum for a small cell base station. Different indoor scenarios with small cells belonging to two different operators have been studied. The results show that the proposed approach is capable to achieve a performance between 96% and 99% of the optimum throughput.

In [27], a Q-learning mechanism for advanced learning of the activity within an unlicensed band is proposed. This mechanism results in enhanced coexistence between LTE LAA and Wi-Fi. Furthermore, the coexistence is further enhanced through a double Q-learning method. This method takes into account both transmit power control of LTE and discontinuous transmission. Simulation results show that the proposed methods are capable to improve both LTE and Wi-Fi performance.

C. ENHANCEMENT OF mLTE-U SCHEME WITH Q-LEARNING

Although 3GPP published the LTE LAA standard that describes a CCA procedure that must be performed before a DL transmission, the ratio between LTE LAA and Wi-Fi TXOP is not balanced, especially when Wi-Fi does not use or support frame aggregation. In order to balance the TXOP of LTE and Wi-Fi, in our previous work [9], we proposed an adaptive LTE LBT scheme named mLTE-U. Similar to LTE LAA, this scheme uses an anchor channel in licensed band together with a secondary channel in unlicensed spectrum, which can be exploited by the eNB to transmit DL traffic. mLTE-U requires a CCA procedure before a DL transmission in the unlicensed spectrum and uses adaptable LTE TXOP followed by an adaptable muting period. The muting period can be exploited by other co-located technologies, such as Wi-Fi, to gain access to the medium. The provided coexistence performance depends on the selection of TXOP and muting period duration. This article further extends our previous work by introducing a Q-learning technique for autonomous selection of the optimal TXOP and muting period by an mLTE-U eNB that can enable fair coexistence between mLTE-U and Wi-Fi. Additionally, this article provides a system model analysis of the mLTE-U scheme in coexistence with Wi-Fi, in comparison to [9], where the mLTE-U scheme has been implemented and evaluated using the NS3 simulation platform.

III. PROBLEM DEFINITION AND THE PROPOSED SOLUTION

Recently, 3GPP published the LTE LAA standard in order to enable the LTE operation in unlicensed spectrum as part of LTE Release 13. In order to satisfy the regulations in regions where an LBT procedure is mandatory, such as Europe and Japan, LTE LAA defines a CCA procedure that must be performed before a DL LTE transmission in the unlicensed spectrum. Before a transmission, an eNB has to evaluate the availability of the channel. If the channel is busy, then it must defer its transmission and perform an exponential backoff. When the channel is idle, then the eNB starts a transmission burst for a duration that ranges from 2 ms up to 10 ms. The transmission duration is defined by four different channel access priority classes. Table 1 presents the different priority classes as they are defined by the 3GPP LTE LAA standard. According to the standard, the priority classes 3 and 4 use a $T_{m\ cot,p}$ that is equal to 10 ms if the absence of any other co-located technology sharing the same channel can be guaranteed on a long term basis. Otherwise, the LTE transmission duration in unlicensed spectrum is limited to 8 ms.

On the contrary, in traditional Wi-Fi network, the AP or the station (STA) transmits only one packet after the medium is estimated as idle, when frame aggregation is not supported or is not enabled. Such transmission typically lasts for a few hundreds of μs. In various widely used Wi-Fi standards such as 802.11a/g frame aggregation is not supported, but even if it is available (e.g. 802.11n/ac [28]), in several cases it is not used depending on the traffic type constraints such as low latency [29]. Additionally, 802.11e uses Enhanced Distributed Channel Access (EDCA) that defines four Access Categories (AC) [7]. Two of these AC, named Background (AC_BK) and Best Effort (AC_BE), define TXOPs of only a single frame. The other two, named Video (AC_VI) and Voice (AC_VO), define TXOPs of 3.008 *ms* and 1.504 *ms* duration respectively. However, these TXOPs are not balanced compared to the TXOPs defined for LTE LAA that can go up to 10 *ms* and although they have defined by the standard, practical implementations rarely use them.

It is clear that the ratio between the transmission duration of LTE and Wi-Fi in the unlicensed spectrum is not balanced as the TXOP duration of LTE LAA is significantly longer compared to the single packet transmission of Wi-Fi. In order to deal with this concern, in our previous work [9], we proposed the mLTE-U coexistence mechanism. mLTE-U is a novel and adaptable technique that enables fair coexistence between LTE and Wi-Fi. Before a transmission in the unlicensed band, mLTE-U must perform an LBT Cat 4 procedure. If the medium is estimated as idle, LTE can transmit DL traffic for a variable TXOP duration, followed by a variable muting period. Without loose of generality, the TXOP is selected in a range of 2 ms up to 20 ms and the muting period is selected in a range of 0 ms up to 20 ms. Fig. 1 shows the mLTE-U scheme.

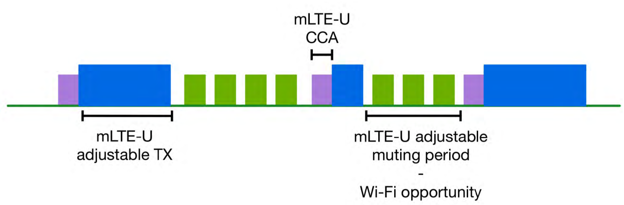

FIGURE 1. **The design of the mLTE-U scheme.**

In [9], the proposed scheme has been evaluated under different coexistence scenarios (low to high LTE and Wi-Fi density), investigating the different combination of TXOP and muting period. This article goes a step further by analytically studying the system model of mLTE-U in coexistence with Wi-Fi and by employing a reinforcement learning technique, more specifically a Q-learning technique, so that an eNB can automatically and autonomously select the optimal configuration parameters (TXOP and muting period) that can lead to fair coexistence with other co-located networks.

IV. SYSTEM MODEL

This section aims to analyze the system model of the proposed mLTE-U scheme, when it coexists with Wi-Fi. All the participating networks operate autonomously and cannot exchange messages with each other. In this work and similar to LTE Release 13, the eNB is able to transmit in the unlicensed spectrum, while the UL traffic is transmitted via the primary licensed band. We consider as active any mLTE-U eNB, Wi-Fi AP and Wi-Fi STA node that has traffic to transmit in unlicensed spectrum. All the active nodes use the same LBT algorithm with random backoff and variable size of CW (similar to LBT Cat 4). For instance, we consider a scenario where one mLTE-U network consisting of one eNB and one UE coexists with one Wi-Fi network consisting of one AP and one STA. If the eNB, the AP and the STA have data to transmit, then all these three nodes are indicated as active. On the other hand, if only the eNB and the AP have data to transmit, then only these two nodes are indicated as active. It is assumed that all the co-located networks transmit in a single unlicensed channel. For the sake of simplicity, we assume that all the networks are in the proximity of each other. This means that every transmission can be determined by the Energy Detection (ED) mechanism of CCA for both mLTE-U and Wi-Fi networks. ED is a function used by CCA to determine the state of the channel, when the received signal cannot be decoded. The CCA mechanism of 802.11 uses also a second function, named Carrier Sense (CS). CS is used when the receiver is able to detect and decode a received Wi-Fi preamble [7].

Both mLTE-U and Wi-Fi use a Carrier Sensing Multiple Access with Collision Avoidance (CSMA/CA) mechanism to compete for the channel access. Before a transmission, every network has to perform CCA in order to sense the channel and discover if it is idle or busy. Before a new transmission or after a successful transmission, a node has to postpone its transmission for Distributed Coordination Function (DCF) Inter-Frame Space (DIFS) plus a random backoff time. The backoff time corresponds to the number of idle timeslots (*ts*) that a node has to sense before a transmission. The number of the *ts* is indicated by the backoff counter, which is randomly selected within the range of the CW. If a transmission is not successful and an acknowledgment (ACK) is not received, the CW increases exponentially. For both mLTE and Wi-Fi the CW ranges from CW_{min} to CW_{max}.

We denote the number of the active mLTE-U eNBs as L and the number of active Wi-Fi APs and active Wi-Fi STAs as A and S respectively. The total number of the active Wi-Fi nodes is denoted as W, where $W = A + S$. The probability that a node tries to transmit at any moment is independent of the previous transmissions. Furthermore, the transmission probability is related to the size of the CW. By assuming that the probability of a transmission to be involved in a collision is very small, the transmission probability of the i-th mLTE-U eNB p_i and the transmission probability of the j-th Wi-Fi node r_j both depend on the CW_{min} and respectively are equal to:

$$p_i = \frac{1}{CW_{\min,i} + 1}, \quad i = 1, \dots, L \tag{1}$$

and

$$r_j = \frac{1}{CW_{\min,j} + 1}, \quad j = 1, \dots, W \tag{2}$$

As in the current model an mLTE-U eNB and a Wi-Fi node use the same CW_{min} value, they have equal probabilities to access the medium.

According to the CCA mechanism that is used by both networks, the time frame can be divided into four different slots:

1) Collision slot T_{col}, meaning that more than one of the co-located nodes (eNBs, APs or STAs) attempt to transmit simultaneously
2) Empty slot T_{empty}, meaning that none of the nodes attempts to transmit
3) Successful mLTE-U transmission slot $T_{mLTE\text{-}U}$, meaning that only one eNB transmits, while the rest eNBs and all the Wi-Fi nodes remain silent
4) Successful Wi-Fi transmission slot $T_{Wi\text{-}Fi}$, meaning that only one Wi-Fi node transmits, while the rest Wi-Fi nodes and all the eNBs remain silent.

Fig. 2 illustrates the system model of mLTE-U when it coexists with Wi-Fi.

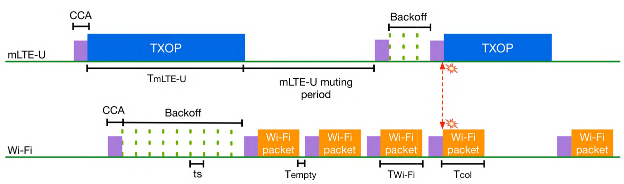

FIGURE 2. The system model of mLTE-U in coexistence with Wi-Fi.

The transmissions of each co-located network are independent and identically distributed (i.i.d.). Hence, the probability that the i-th mLTE-U eNB transmits successfully during a slot is:

$$p_{succ,i}^{\text{mLTE-U}} = p_i \times \prod_{l \neq i}^{L} (1 - p_l) \times \prod_{j=1}^{W} (1 - r_j) \tag{3}$$

Similarly, the probability that the j-th Wi-Fi node transmits successfully during a slot is:

$$p_{succ,j}^{Wi-Fi} = r_j \times \prod_{w \neq j}^{W} (1 - r_w) \times \prod_{i=1}^{L} (1 - p_i) \tag{4}$$

The probability that a slot is empty is expressed as:

$$p_{\text{empty}} = \prod_{i=1}^{L} (1 - p_i) \times \prod_{j=1}^{W} (1 - r_j) \tag{5}$$

while the probability that a collision occurs in a slot is given by:

$$p_{\text{col}} = 1 - p_{\text{empty}} - \sum_{i=1}^{L} (p_{succ,i}^{\text{mLTE-U}}) - \sum_{j=1}^{W} (p_{succ,j}^{Wi-Fi}) \tag{6}$$

The total duration of the slots is expressed as:

$$T_{\text{total}} = Tot_{\text{empty}} + Tot_{\text{col}} + Tot_{\text{Wi-Fi}} + Tot_{\text{mLTE-U}} \tag{7}$$

where Tot_{empty} and Tot_{col} denote the total duration of the empty and the collision slots respectively, $Tot_{Wi\text{-}Fi}$ denotes the total duration of the successful Wi-Fi transmissions and $Tot_{\,mLTE\text{-}U}$ represents the total duration of the successful mLTE-U transmissions in unlicensed spectrum.

Furthermore, the total combined throughput of Wi-Fi can be calculated by:

$$Thr_{\text{Wi-Fi}} = \sum_{j=1}^{W} \left(\frac{D_{Thr,j}^{Wi-Fi}}{T_{\text{total}}}\right) \tag{8}$$

where $D_{Thr,j}^{Wi-Fi}$ is the transmitted payload of Wi-Fi node j. Similarly, the total combined throughput of mLTE-U in the unlicensed band is expressed as:

$$Thr_{\text{mLTE-U}} = \sum_{i=1}^{L} \left(\frac{D_{Thr,i}^{\text{mLTE-U}}}{T_{\text{total}}}\right) \tag{9}$$

where $D_{Thr,i}^{\text{mLTE-U}}$ is the transmitted payload of the i-th mLTE-U eNB.

A. RESERVATION SIGNAL

An mLTE-U eNB must perform a CCA procedure before a transmission to estimate if the channel is idle or not. Hence, the medium can be sensed as idle at any time. However, LTE is a scheduled technology on a sub-frame level, meaning that every 1 ms the eNB scheduler assigns the wireless resources to the active UE. This means that every data transmission starts at the beginning of a subframe. To deal with this issue and similar to our previous work in [9], a reservation signal is used for mLTE-U in order to reserve the channel after it is sensed as idle and before the beginning of the next subframe. Fig. 3 illustrates the use of the reservation signal.

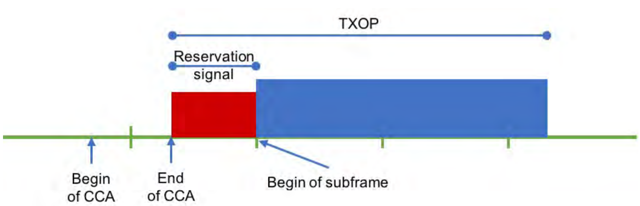

FIGURE 3. The reservation signal of the mLTE-U scheme.

The reservation signal is modeled by a uniformly distributed random variable in the interval [0,1]. A value close to zero corresponds to a short duration of reservation signal. This means that the channel is sensed as idle towards the ending of a subframe. A value close to one means that the channel is sensed as idle in the beginning of a subframe. Thus, the reservation signal is transmitted for the rest of the subframe duration. The duration of the reservation signal is deducted from the TXOP duration of the mLTE-U scheme. For this reason, the minimum examined TXOP duration is 2 ms.

V. FAIR COEXISTENCE

This section discusses the way that the two different parameters of mLTE-U scheme, named TXOP and muting period, can be selected in order to ensure fair coexistence between co-located mLTE-U and Wi-Fi networks. A fair coexistence scheme should be able to provide to all the active nodes in the unlicensed spectrum equal opportunities to the wireless resources. This must be done in a technology-agnostic way, as all the nodes must be treated equally. According to this approach, all the active mLTE-U eNBs, Wi-Fi APs and Wi-Fi STAs should be able to gain equal spectrum access.

In an ideal world in which the different wireless technologies can communicate with each other, exchange their spectral requirements and operate altruistically, the distribution of the wireless resources could be done in a fair and harmonious way. However, in the real wireless world, several diverse wireless technologies that have been designed, each having different target group, different principles and different requirements are forced to coexist with each other. Additionally, the channel access mechanism of the different technologies vary significantly between each other. In [9], we saw that the obtained throughput, as well as the percentage of channel occupancy are good indicators for measuring the fairness that a coexistence technique can provide. According to this approach, the parameters of mLTE-U must be selected in a way that every co-located network can achieve an equal ratio of throughput, compared to the maximum throughput that it can achieve when it operates in standalone mode, meaning that it operates without any other co-located network.

This assumption requires that every node is able to identify potential co-located networks and aproximate the number of transmitting devices. This can be achieved using a wireless technology recognition technique. Recently, the technology recognition problem has attracted the attention of the wireless community. As result, several techniques (e.g. [8] and [30]) have been proposed and can be used by an mLTE-U network to identify the amount and the type of co-located wireless technologies. Based on this information, an mLTE-U network can select the TXOP and muting period so that it can offer the desired proportional fair throughput. Further discussion on the nature of these techniques is not in the scope of this article and it is assumed that such a technology recognition technique is available to an mLTE-U eNB.

In our system, the target throughput of an mLTE-U network can be expressed as:

$$Thr^{\text{mLTE-U}}_{target,i} = \frac{Thr^{\text{mLTE-U}}_{standalone,i}}{L+W} \tag{10}$$

where $Thr^{\text{mLTE-U}}_{standalone,i}$ is the throughput that the mLTE-U network i can achieve in standalone operation using the maximum TXOP configuration (20 ms) and a muting period that is equal to zero. A muting period that is equal to zero ensures that the eNB can start competing for the medium immediately after finishing a transmission of TXOP duration. Moreover, the highest TXOP ensures that the eNB can transmit for a longer period without interruption. The configuration of TXOP has an impact on the obtained throughput. For a lower TXOP, the eNB has to perform a CCA procedure more frequently compared to a higher TXOP. This forces the eNB to spend more time evaluating the channel compared to the case in which it uses a high TXOP.

Considering the system that is described in Section IV, the configuration of TXOP and muting period for an mLTE-U eNB must be selected according to the following optimization problem:

$$\begin{aligned}
&(TXOP^*_i, muting^*_i) \\
&\quad = arg \max_{TXOP,muting} (|\,(|\,Thr^{\text{mLTE-U}}_{target,i} \\
&\qquad\qquad -Thr^{\text{mLTE-U}}_{i}\,|) - Thr^{\text{mLTE-U}}_{target,i}\,|) \\
&\quad \text{s.t. } C1:\ 0 \le p_{\text{i}} \le 1, \quad i = 1, \ldots, L \\
&\qquad C2:\ 0 \le r_{\text{j}} \le 1, \quad j = 1, \ldots, W \\
&\qquad C3:\ |\,Thr^{\text{mLTE-U}}_{target,i} - Thr^{\text{mLTE-U}}_{i}| \le \zeta, \quad i = 1, \ldots, L \\
&\qquad C4:\ TXOP \in [TXOP_{min}, TXOP_{max}] \\
&\qquad C5:\ muting \in [muting_{min}, muting_{max}]
\end{aligned} \tag{11}$$

This problem guarantees that the optimal TXOP and muting period values will be selected so that the obtained mLTE-U throughput will be maintained close to the target value, offering this way fair coexistence with other co-located mLTE-U or Wi-Fi networks. The first constraint (C1) refers to the transmission probability of an mLTE-U eNB, while the second constraint (C2) refers to the transmission probability of a Wi-Fi node. The third constraint (C3) indicates that the absolute difference between the target throughput of eNB i and the throughput that eNB i achieves after the TXOP and muting period adjustment remains within a tolerance range that is defined by ζ. This constraint ensures that the mLTE-U throughput will remain in an acceptable range close to the target throughput, giving transmission opportunities to other co-located networks. The fourth (C4) and the fifth (C5) constraints ensure that the selected values of TXOP and muting period will be within an acceptable range.

VI. PROPOSED Q-LEARNING FOR FAIR COEXISTENCE BETWEEN MLTE-U AND WI-FI

This section discusses how Q-learning can be used in the described model so that an eNB of an mLTE-U network can learn from the environment and autonomously select the appropriate TXOP and muting period combination that can enable fair coexistence with other co-located mLTE-U or Wi-Fi networks.

Q-learning is a type of Reinforcement Learning (RL) in the area of machine learning. According to Q-learning, an agent in a state s selects and performs an action a. After the action a, it observes the environment and receives a reward r for this specific action a. A discount factor γ models the percentage that future rewards are taken into account compared to immediate rewards. Hence, the scope of Q-learning is to find the optimal policy π^* for selecting an action in a given state that maximizes the value of the total reward. In order to learn this policy an agent has to estimate a value-function through experience. This function is called Q-function $Q^{\pi}(s, a)$ [31]. The Q-function expresses the expected accumulated discounted future reward r that is obtained at time t by selecting an action a in a state s and by following thereafter a policy π. This can be expressed as follows:

$$Q^{\pi}(s, a) = E(\sum_{t=1}^{\infty} \gamma^{t-1} r_t | s_1 = s, a_1 = a, \pi) \tag{12}$$

Q-learning does not require a prior environment model and it can be applied to any given Markov Decision Process (MDP) model. The interaction of an agent with the dynamic stochastic environment is represented by an experience tuple (s_t, a_t, s_{t+1}, r_t), where s_t is the state of an agent at time t and a_t is the action that the agent chooses at time t from the set of the available actions. Then, the agent moves to a new state s_{t+1} at time $t+1$, in which a reward r_t associated with the transition from the state s_t to the state s_{t+1} is determined. The Q-learning process can be represented by the following update equation:

$$Q_{t+1}(s_t, a_t) \leftarrow Q_t(s_t, a_t) + \eta[r_t + \gamma Q' - Q_t(s_t, a_t)] \tag{13}$$

where η is the learning rate and γ is the discount factor. The learning rate can be set between 0 and 1. It determines the percentage that the newly learned information will overwrite the older knowledge. By setting the learning rate to 0 the Q-values are never updated and as result nothing is learned. By setting it to a high value such as 0.9 means that the agent learns at a faster rate. The discount factor γ takes values in the range [0,1]. When it is set to a value closer to one, the agent will consider future rewards with greater weight. The value of Q' indicates the maximum reward that can be attained in a state following the current one. In other words, it expresses the reward for performing the optimal action from the current state and is denoted as follows:

$$Q' = \max_{a \in A} Q_t(s_{t+1}, a_t) \tag{14}$$

where A is the set of all the possible actions ($A = \{a_1, a_2, \ldots, a_i\}$) of the i-th agent.

A. DEFINITION OF Q-LEARNING ELEMENTS

In the investigating learning scenario, an eNB of an mLTE-U network must learn to be configured with the appropriate TXOP and muting period values that offer fair coexistence with other mLTE-U or Wi-Fi networks using Q-learning. To this end, the agents, states, actions and rewards for the Q-learning algorithm are defined as follows:

1) AGENT

In the investigated multi-agent scenario, every i-th eNB of an mLTE-U network is an agent, $\forall i = 1, \ldots, L$.

2) STATE

For every agent the state is selected by the interaction with the environment. The state s_t^i for an agent i at the time instance t is represented as $s_t^i = \{TXOP^i, muting^i\}$, where $TXOP^i \in [2, 20]$ and $muting^i \in [0, 20]$ is the TXOP and the muting period for the agent i respectively.

3) ACTION

The action of the agent i is to select the TXOP and muting period that can offer fair coexistence with other co-located wireless technologies.

4) REWARD

The reward for an action a of the agent i is given by the following function:

$$r_i^{\text{mLTE-U}} = \begin{cases} \beta \times (| (| Thr_{target,i}^{\text{mLTE-U}} - Thr_i^{\text{mLTE-U}} |) \\ \qquad\qquad -Thr_{target,i}^{\text{mLTE-U}} |) \\ \text{for perf\_dif} < \zeta \\ -100 \\ \text{for perf\_dif} \geq \zeta \end{cases} \tag{15}$$

where β determines the fraction of the positive reward, perf_dif $=| Thr_{target,i}^{\text{mLTE-U}} - Thr_i^{\text{mLTE-U}} |$ is the absolute value of the difference in performance between the target throughput of i-th eNB and the throughput that the i-th eNB achieves after action a has been performed. Similar to the third constraint in (11), ζ defines a tolerance range for the achieved throughput in a state s. Hence, if after an action the obtained throughput is close to the target throughput ($Thr_{target,i}^{\text{mLTE-U}}$) meaning that their absolute difference is within the tolerance range, then the agent receives a reward that is proportional to the deviation of the obtained throughput from the target throughput. Otherwise, the agent receives a negative reward.

B. EXPLORATION STRATEGY

The scope of Q-learning is to find an optimal strategy in the selection of an action a from a state s. Hence, a balance between exploration and exploitation must be found. When an agent exploits, it selects the currently expected optimal action (Q'). On the other hand, when it explores, it selects randomly an action in the hope that it will offer a higher cumulative reward in the future. Hence, by exploring, an agent investigates new actions, while by exploiting it selects the optimal action from the already investigated actions. In this article, the ϵ-greedy policy is used as exploration strategy. ϵ-greedy uses $0 \leq \epsilon \leq 1$ in order to decide if the agent will explore or exploit in every step. The agent chooses a random action (explore) with probability ϵ and the action with the highest Q-value from the current state (exploit) with probability $1 - \epsilon$. When ϵ is configured with a high value, more exploration actions are selected by the agent. This is useful for an agent to learn the environment and the optimal policy.

In this article, an adjustable policy for the value of ϵ is used. Initially or every time that a change to the wireless environment is sensed by the technology recognition technique, ϵ will be set to a high value (e.g. 1) in order to quickly explore different states. After a number of iterations i_ϵ the value of ϵ will be reduced by a p_ϵ value (e.g. 0.05), until a minimum value of ϵ (m_ϵ) is reached (e.g. 0.05) or until the Q-learning converges to the optimal solution.

Algorithm 1 presents the proposed Q-learning procedure as it is described above and is required by an independent mLTE-U network to select an optimal configuration that enables fair coexistence with the co-located LTE or Wi-Fi networks.

Regarding the computational complexity of the Q-learning mechanism and similarly to other learning methods, a learning phase is required. During this phase, an agent discovers the environment by investigating different possible actions in every possible state. However, once the environment is learned, the best action can be performed in any given state resulting in the optimal solution. In case that the technology recognition technique is not completely accurate, then the proposed scheme can still achieve performance close to the optimal one.

VII. SIMULATION ENVIRONMENT

In order to evaluate the proposed mLTE-U scheme and the Q-learning algorithm for optimal and autonomous selection of the mLTE-U parameters, simulations have been performed using MATLAB.

For an mLTE-U network only the throughput in the unlicensed spectrum is taken into consideration. Furthermore, it is assumed that only LTE DL data traffic is transmitted in the unlicensed spectrum, while the LTE UL traffic, the LTE control signals and the Hybrid Automatic Repeat Request (HARQ) are maintained in the licensed band of the operator.

Regarding the Wi-Fi network, 802.11n mode has been selected for the simulation model. This mode allows operation in 5 GHz unlicensed band. Additionally, it is assumed that frame aggregation is disabled, so that only a single packet is transmitted after the channel is estimated as idle. Table 2 presents the system parameters that have been used for Wi-Fi.

The average backoff time for a Wi-Fi transmission can be expressed as:

$$T_{\text{Av\_BO}} = CW_{\min} \times \frac{ts}{2} \tag{16}$$

Additionally, the duration of the acknowledgment is given by:

$$T_{\text{ack}} = T_{\text{plcp}} + \left\lceil \frac{L_s + L_{\text{ack}} + L_t}{n_{\text{sym}}} \right\rceil \times T_{\text{sym}} \tag{17}$$

The duration (T_{plcp}) of Physical Layer Conformance Procedure (PLCP) is $20\mu s$ and corresponds to $8\mu s$ for the Short

Algorithm 1 Q-learning for mLTE-U optimal configuration selection

Initialization:
$TXOP_{min}$, set minimum TXOP value
$TXOP_{max}$, set maximum TXOP value
$muting_{min}$, set minimum muting value
$muting_{max}$, set maximum muting value
t_r, technology recognition result
ϵ, set the ϵ-greedy to a high value (e.g. 1)
i_ϵ, set the number of the iterations before reduce ϵ
p_ϵ, set the rate in which ϵ will be reduced
m_ϵ, set the minimum value of ϵ
ζ, set the throughput tolerance
β, set the fraction of the positive rewards
η, set the learning rate
γ, set the discount factor
for *every i-th mLTE-U eNB, where i = 1, ..., L* **do**
  Set $iteration = 0$, $Q_{i,0}(s, a) = 0$
  Randomly choose a starting state
  $s_{i,0} = TXOP_{i,0}, muting_{i,0}$ and evaluate it
end

Learning procedure:
while *(t_r has not changed) OR (convergence is not achieved)* **do**
  if *(a number of iterations i_ϵ has been reached) &* $(\epsilon > m\_\epsilon)$ **then**
    $\epsilon = \epsilon - p\_\epsilon$
  end
  Randomly choose $prob\_e \in [0, 1]$
  if $prob\_e < \epsilon$ **then**
    [exploration procedure]
    Select the next action $a_{i,t}$ randomly
  else
    [exploitation procedure]
    Select the next action $a_{i,t}$ based on the max(Q-value): $maxQ_{i,t}(s_{i,t}, a_{i,t})$
  end
  Execute $a_{i,t}$
  Receive an immediate throughput $Thr_{i,t}^{\text{mLTE-U}}$
  if $(|\ Thr_{target,i}^{mLTE\text{-}U} - Thr_{i,t}^{mLTE\text{-}U}\ | < \zeta)$ **then**
    $r_{i,t}^{\text{mLTE-U}} = \beta \times (|\ (|\ Thr_{target,i}^{\text{mLTE-U}} - Thr_{i,t}^{\text{mLTE-U}}\ |) - Thr_{target,i}^{\text{mLTE-U}}\ |)$
  else
    $r_{i,t}^{\text{mLTE-U}} = -100$
  end
  Update the Q-table (according to 13) as follows:
  $Q_{\text{i,t+1}}(s_{i,t}, a_{i,t}) \leftarrow Q_{\text{i,t}}(s_{i,t}, a_{i,t}) + \eta[r_{i,t}^{\text{mLTE-U}} + \gamma \max_{a_{i,t} \in A} Q_{\text{i,t}}(s_{i,t+1}, a_{i,t}) - Q_{\text{i,t}}(s_{i,t}, a_{i,t})]$
  Next state: $s_{i,t+1}$
end

Monitor the wireless environment:
while *(true)* **do**
  Periodically monitor the wireless environment
  if *(a change is identified)* **then**
    Update t_r
    Restart **Learning procedure**
  end
end

TABLE 2. Wi-Fi simulation parameters.

Parameter	Value
Wi-Fi mode	802.11n
Frame aggregation	no
Bandwidth	20 MHz
DIFS duration	34 μs
SIFS duration	16 μs
Timeslot duration (ts)	9 μs
PLCP preamble + Headers Duration (T_{plcp})	20 μs
PLCP service field (L_s)	16 bits
MAC header (L_{MAC_h})	224 bits
Tail bits (L_t)	6 bits
ACK length (L_{ack})	112 bits
Payload (D)	12000 bits
OFDM Symbol duration (T_{sym})	4 μs
Number of bits per OFDM symbol (n_{sym})	216 bits
CW_{min}	15
CW_{max}	1023
RTS/CTS	no

Training Field (STF), $8\mu s$ for the Long Training Field (STF) and $4\mu s$ for the SIGNAL field.

The duration of a data-packet transmission is given by:

$$T_{\text{data}} = T_{\text{plcp}} + \left\lceil \frac{L_s + L_{\text{MAC\_h}} + D + L_t}{n_{\text{ sym}}} \right\rceil \times T_{\text{sym}} \quad (18)$$

Hence, the total duration of a successful Wi-Fi transmission can be expressed as:

$$T_{\text{suc}} = T_{\text{DIFS}} + T_{\text{Av\_BO}} + T_{\text{SIFS}} + T_{\text{ack}} + T_{\text{data}} \quad (19)$$

For both mLTE-U and Wi-Fi networks 20 MHz of bandwidth is used. For Wi-Fi, 64-Quadrature Amplitude Modulation (QAM) modulation scheme and 3/4 coding scheme has been used that correspond to the 6th Modulation and Coding Scheme (MCS) Index [7]. On the other hand, for mLTE-U transmission in the unlicensed spectrum, the transmission data rate is equal to 150 Mbps. This corresponds to 2×2 MIMO, 64-QAM, 28th MCS Index and 26th Transport Block Size (TBS) Index, as it is defined in 3GPP specs 36.213 [5].

During the simulation, it is assumed that all the nodes for both mLTE-U and Wi-Fi networks are in the proximity of each other. This way, during every transmission the ED threshold is surpassed and the backoff mechanisms of mLTE-U and Wi-Fi are triggered. The ED threshold of the mLTE-U CCA mechanism is equal to the ED threshold of Wi-Fi.

Concerning the Q-learning parameters, they are listed in Table 3. The ϵ parameter initially takes a high value (e.g. 1) in order to explore fast new states. As the number of iterations increases and all or most of the states are reached at least once, the ϵ value decreases by p_ϵ, until a minimum value of ϵ is reached (m_ϵ). During the simulations, the number of

TABLE 3. Q-learning simulation parameters.

Parameter	Value
range of ϵ value	$[1 - 0.05]$
learning rate (η) value	0.7
discount factor (γ) value	0.9
rate of ϵ reduction (p_ϵ)	0.05
minimum value of ϵ (m_ϵ)	0.05
number of iterations before reduction of ϵ (i_ϵ)	399
throughput tolerance (ζ) value (Mbps)	3
fraction of positive rewards (β) value	0.2
maximum iteration number	10000

iterations before ϵ decreases is computed as:

$$i\_\epsilon = (TXOP_{max} - TXOP_{min} + 1)(muting_{max} - muting_{min} + 1) \quad (20)$$

that corresponds to the total number of the possible states.

VIII. PERFORMANCE EVALUATION

A. STANDALONE OPERATION FOR mLTE-U AND Wi-Fi

This section presents the performance of the designed system, when mLTE-U and Wi-Fi operate in standalone mode. Thus, they do not need to compete for the wireless medium with other co-located networks. Both mLTE-U and Wi-Fi networks consist of one base station and one end-device.

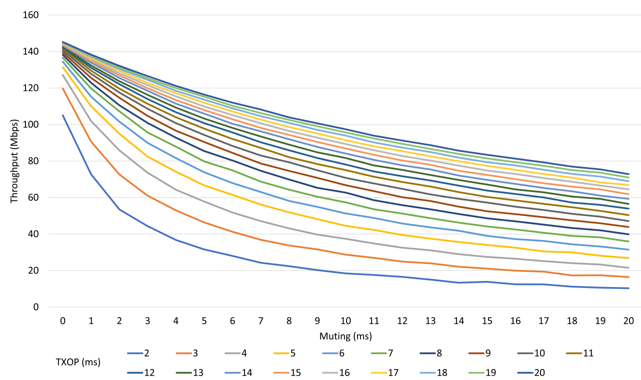

FIGURE 4. Throughput of mLTE-U for the different TXOP and muting period configurations, during the standalone scenario.

Fig. 4 illustrates the obtained DL throughput results of mLTE-U network in standalone mode. The x-axis holds the different muting period configurations in ms ranging from 0 *ms* to 20 *ms*. The different TXOP durations in ms ranging from 2 *ms* to 20 *ms* are representing with different colors. Finally, the y-axis shows the obtained throughput in Mbps for every possible combination of TXOP and muting period.

From the figure, it is clear that the throughput for every different TXOP decreases as the duration of the muting period increases. Of course, this is to be expected as a higher muting period increases the idle period of an eNB. Respectively, it can be seen that for a specific muting period, the obtained throughput increases as the TXOP increases. As the TXOP duration increases, the mLTE-U has to perform less often a CCA procedure before it transmits again. This has an impact on the obtained throughput, as for higher TXOP the eNB spends less time evaluating the channel compared to a scenario in which a lower TXOP duration is used.

Hence, the minimum obtained throughput corresponds to an mLTE-U configuration, in which TXOP has the smallest value (2 *ms*) and it is followed by a muting period of the longest duration (20 *ms*). On the contrary, the maximum obtained throughput can be achieved when the maximum TXOP is used (20 *ms*) followed by the minimum muting period (0 *ms*).

According to the simulation results and after the introduction of CCA, the highest throughput value of mLTE-U for $TXOP = 20\ ms$ and $muting = 0\ ms$ is 145.28 *Mbps*. This value will be used for the computation of the target mLTE-U throughput in (10) that is used by the Q-learning algorithm. Regarding the Wi-Fi network, the obtained standalone throughput is stable over time and corresponds to 30.8 *Mbps*.

B. mLTE-U AND Wi-Fi COEXISTENCE

In this section, coexistence scenarios between mLTE-U and Wi-Fi of high interest are discussed. This will help the reader to understand the role of Q-learning in selecting the mLTE-U configurations that can offer fair coexistence with other co-located networks. Further details on the coexistence between mLTE-U and Wi-Fi can be found in [9].

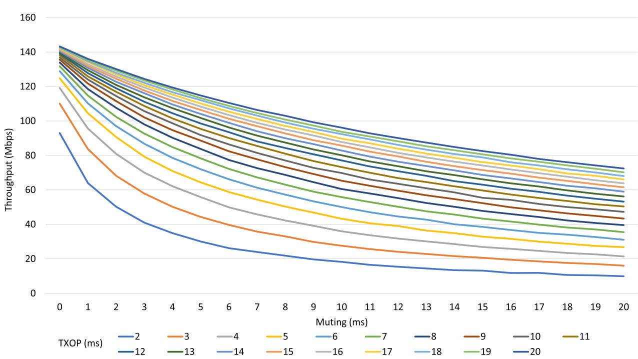

FIGURE 5. **Throughput of mLTE-U during the single mLTE-U and single Wi-Fi coexistence scenario.**

1) EVALUATION OF SINGLE mLTE-U AND SINGLE Wi-Fi COEXISTENCE

In this scenario, one mLTE-U network coexists with one Wi-Fi network. The mLTE-U network consists of one eNB and one UE, while the Wi-Fi network consists of one AP and one STA. Both networks transmit only DL traffic. Fig. 5 depicts the mLTE-U throughput and Fig. 6 the Wi-Fi throughput for every possible combination of TXOP and muting period. In both figures, the x-axis holds the different muting period configurations in ms. The different TXOP configurations (in ms) are depicted with different colors. The y-axis

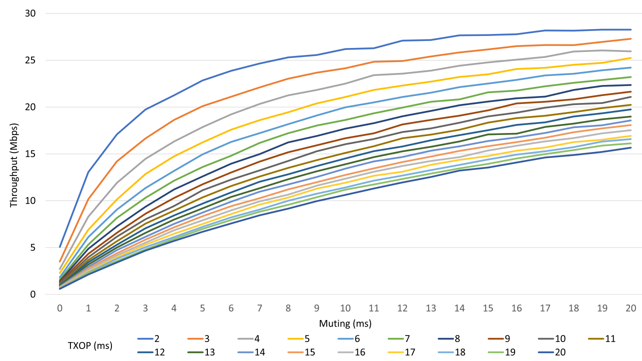

FIGURE 6. **Throughput of Wi-Fi during the single mLTE-U and single Wi-Fi coexistence scenario.**

presents the obtained throughput in Mbps for every combination of TXOP and muting period.

As it can be observed and similar to the standalone scenario, the mLTE-U throughput increases as the TXOP increases. Also, a shorter muting period offers higher throughput compared to a longer one as mLTE-U can compete more often for accessing the medium. Furthermore, the throughput values are slightly lower compared to the standalone scenario. This occurs due to the co-located Wi-Fi network that competes for the medium and eventually gains access to it. On the other hand, the Wi-Fi throughput increases when the muting period of mLTE-U increases. This is to be expected, as Wi-Fi can exploit the muting period for further transmissions. Additionally, the Wi-Fi throughput is inversely proportional to the TXOP of mLTE-U. During a short TXOP, Wi-Fi has more often opportunities to compete for the medium and access it compared to a longer TXOP during which mLTE-U occupies the medium for longer period of time.

2) EVALUATION OF MULTIPLE mLTE-U AND MULTIPLE Wi-Fi COEXISTENCE

In this scenario, multiple mLTE-U and multiple Wi-Fi networks coexist among each other. More specifically, three mLTE-U networks coexist with three Wi-Fi networks creating this way a dense wireless environment. Each one of the mLTE-U and Wi-Fi networks consists of one base station and one end-device. Each network transmits only DL traffic. Similarly to the previous subsection (VIII-B.1), Fig. 7 and Fig. 8 show the mLTE-U combined throughput and the Wi-Fi combined throughput respectively.

Fig. 8 clearly indicates that the performance of the Wi-Fi networks is severely impacted by the co-located mLTE-U networks for most of the mLTE-U configurations. Only when mLTE-U is configured with a short TXOP that is followed by a relatively long muting period, the combined throughput of Wi-Fi is improved. In case of multiple mLTE-U nodes, there is a high possibility that a muting period of an mLTE-U network is exploited by the TXOP of another mLTE-U network. This impact becomes higher when the mLTE-U networks are configured to use a high TXOP duration combined with a low

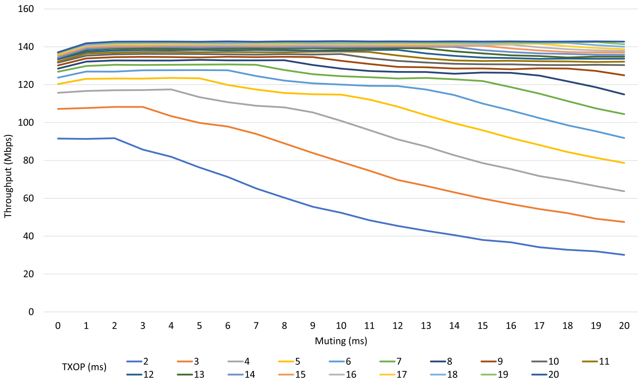

FIGURE 7. **Combined throughput of mLTE-U during the multiple mLTE-U and multiple Wi-Fi coexistence scenario.**

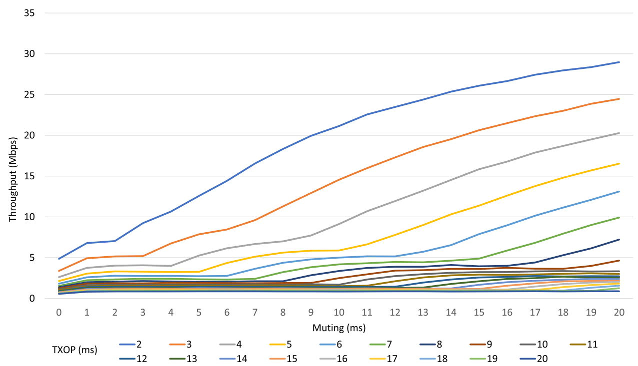

FIGURE 8. **Combined throughput of Wi-Fi during the multiple mLTE-U and multiple Wi-Fi coexistence scenario.**

muting period. However, when the mLTE-U networks use a short TXOP and a high muting period, they remain silent simultaneously for a longer period and Wi-Fi can exploit the remaining muting period in order to transmit. Furthermore, in case of multiple Wi-Fi networks the exploitation of a muting period is less optimal as they compete among each other to access the medium.

C. FAIR COEXISTENCE USING Q-LEARNING

As shown in the previous subsection, the performance of coexisting mLTE-U and Wi-Fi networks depends on the density of the environment, as well as on the configuration of mLTE-U. The numerous combinations of TXOP and muting period offer different coexistence conditions that vary based on the number of co-located networks. As a wireless environment is dynamic and new networks are activated and deactivated often, it is important for a coexistence scheme to be self-adaptive. This section discuses the way that Q-learning technique, as it has been discussed in Section VI, can assist an mLTE-U network in optimally selecting the TXOP and muting period in order to provide fair coexistence with other co-located wireless technologies in unlicensed spectrum.

1) Q-LEARNING FOR SINGLE mLTE-U AND SINGLE Wi-Fi COEXISTENCE

Fig. 9 illustrates the convergence of the Q-learning algorithm during the scenario in which one mLTE-U network coexists

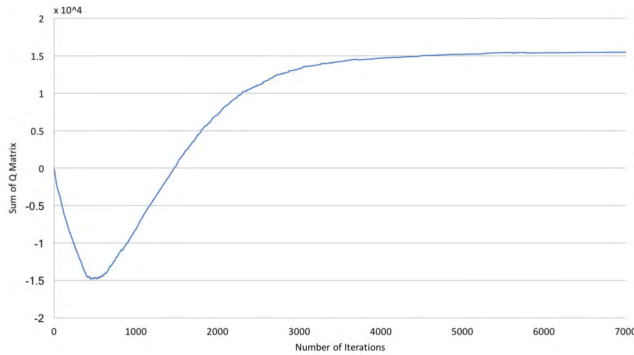

FIGURE 9. **Convergence of Q matrix sum during the learning process for the single mLTE-U and single Wi-Fi scenario.**

with one Wi-Fi network, similar to Section VIII-B.1. On the horizontal axis is the number of iterations and on the vertical axis is the sum of the values in the Q matrix. When the sum of the Q matrix converges, the agent has learned the current environment and can perform the optimal actions in any state.

It can be observed that in the beginning of the learning process the sum of Q matrix decreases. This occurs as initially due to the high degree of exploration, the agent (mLTE-U eNB) tries many different states. Most of these states do not offer the desired fairness. This way the agent receives low rewards. As the learning continues, the agent locates the states that can provide fair coexistence with the Wi-Fi network, increasing the received reward. After a sufficient amount of iterations (e.g. 3000), it can be seen that the agent has learned the configurations that can lead to fair coexistence and the sum of Q matrix starts converging.

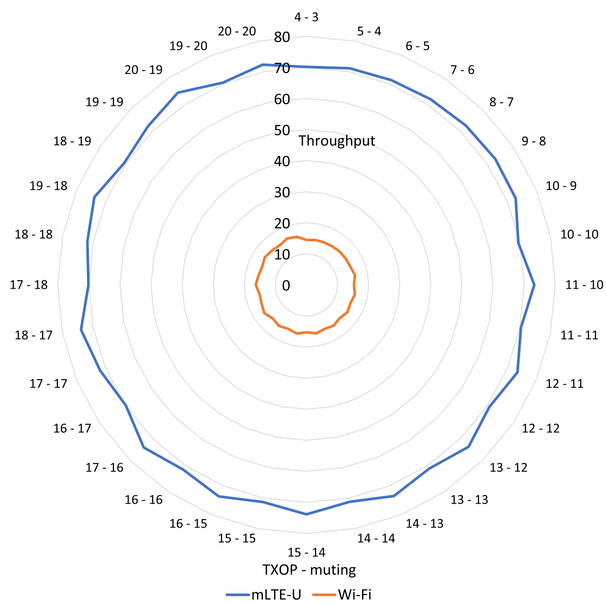

FIGURE 10. **Throughput of mLTE-U and Wi-Fi for the selected by Q-learning configurations of TXOP and muting period during the single mLTE-U and single Wi-Fi scenario.**

Fig. 10 presents the throughput of mLTE-U and Wi-Fi for the selected by Q-learning configurations (TXOP and muting period) and for the same scenario as above, where

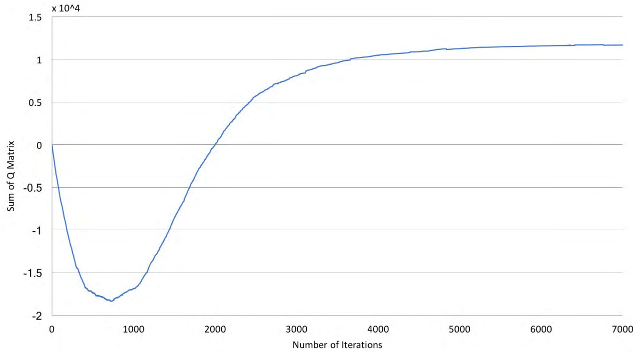

FIGURE 11. Convergence of Q matrix sum during the learning process for the multiple mLTE-U and multiple Wi-Fi scenario.

one mLTE-U network coexists with one Wi-Fi. The TXOP and muting period configurations that have been learned by Q-learning are able to provide to the mLTE-U network a throughput that is in the desired range of $Thr_{target}^{\text{mLTE-U}} \pm \zeta$, where in the specific scenario and from (10) $Thr_{target}^{\text{mLTE-U}} = 72.64$ *Mbps* and $\zeta = 3$ *Mbps*. As can be seen from the results, all the selected configurations are capable to provide the desired fair coexistence with Wi-Fi, as the co-located Wi-Fi network is able to obtain a throughput close to 15 *Mbps*. Hence, both networks can achieve half of the throughput that can be reached during the respective standalone operation.

Based on the traffic requirements that an eNB must satisfy, it can select the appropriate configuration among the ones that have been identified by the Q-learning procedure and can provide fair coexistence with the co-located networks. For instance, in case of voice traffic (AC_VO), an mLTE-U network can select a configuration that requires a shorter muting period. On the other hand, when best effort traffic (AC_BE) must be served, an mLTE-U network can select a configuration that offers a longer muting period combined with a shorter TXOP.

2) Q-LEARNING FOR MULTIPLE mLTE-U AND MULTIPLE Wi-Fi COEXISTENCE

Fig. 11 presents the convergence of the Q-learning algorithm for the coexistence scenario similar to Section VIII-B.2, in which three mLTE-U networks and three Wi-Fi networks coexist with each other.

By observing Fig. 9 and Fig. 11, it can be seen that in case of multiple mLTE-U and Wi-Fi networks (Fig. 11) the sum of the Q matrix initially decreases in a higher grade compared to the case of a single mLTE-U and Wi-Fi network (Fig. 9). In the case of multiple mLTE-U and Wi-Fi networks, many co-located networks have to gain equal access to the medium. Hence, the mLTE-U configurations that can offer fair coexistence are limited compared the configurations of the single mLTE-U and Wi-Fi network. For this reason, during the first iterations of Q-learning, an agent will explore more states that give a negative reward, which entails a reduced sum of Q matrix. As the agent learns the environment and approaches the target, it chooses states that can give high reward, increasing the sum of Q matrix, until it finally converges.

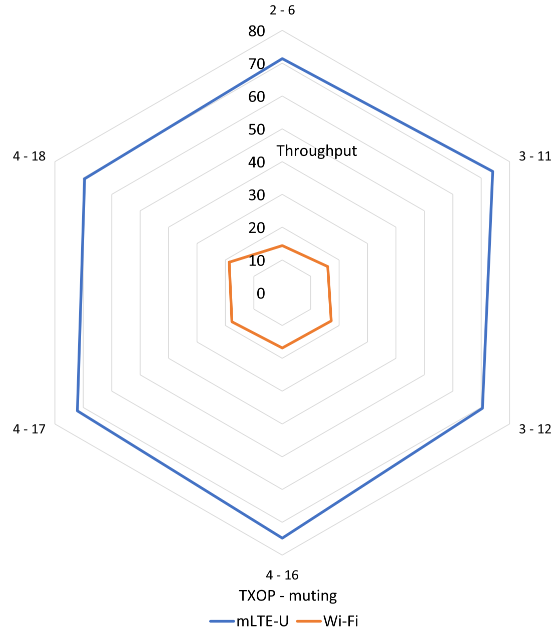

FIGURE 12. Throughput of mLTE-U and Wi-Fi for the selected by Q-learning configurations of TXOP and muting period during the multiple mLTE-U and multiple Wi-Fi scenario.

Fig. 12 illustrates the TXOP and muting period configurations that can offer fair coexistence during this dense scenario, as they have been selected by the Q-learning mechanism. As discussed, it can be observed that compared to the single mLTE-U and single Wi-Fi scenario, the desired combinations are fewer due to the multiple coexisting networks.

3) FURTHER DISCUSSION

Q-learning is fundamentally designed to be able to adapt to the changes of the environment. This way, an agent can update the Q-table and learn new optimal actions towards the achievement of its target. Regarding the mLTE-U scheme, a change in the status of the wireless environment can be identified using a technology recognition scheme. Such change can be the activation of a new network or the deactivation of a previously active network.

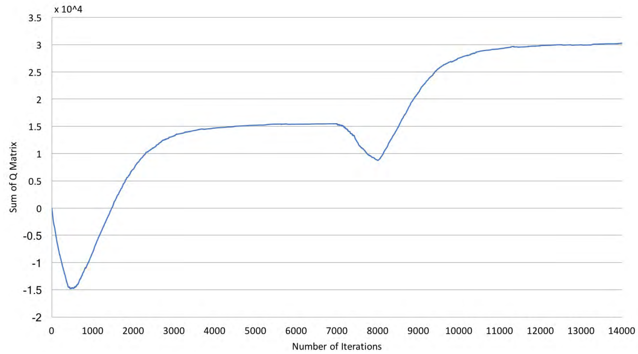

FIGURE 13. Convergence of Q matrix sum during the learning process and adaptation to the changes of the wireless environment.

Fig. 13 shows the convergence of Q-learning for a scenario in which initially one mLTE-U network coexists with one

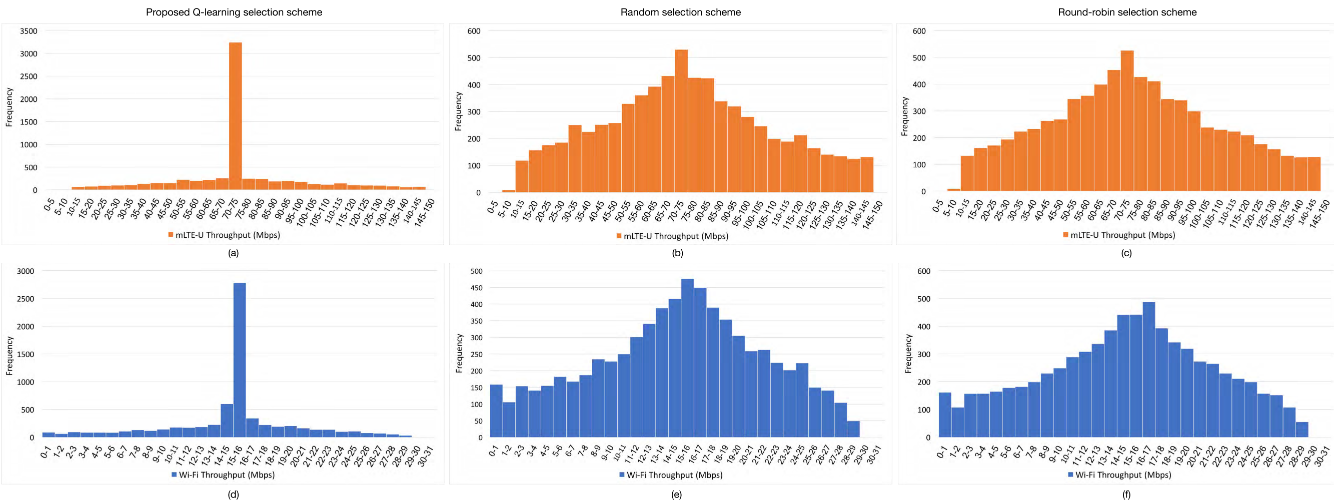

FIGURE 14. Throughput histogram of mLTE-U and Wi-Fi for the proposed Q-learning, random and round-robin selection schemes.

Wi-Fi network and at some point a second mLTE-U network is activated. As it can be seen, the first part of the diagram is similar to the one that is depicted in Fig. 9, as only one mLTE-U network coexists with one Wi-Fi. After the 7000th repetition, a new mLTE-U network is activated. Then, an agent starts identifying the new mLTE-U parameters that can offer fair coexistence regarding the new conditions in the wireless environment using Q-learning. At this point, the ϵ value of the $\epsilon-$greedy exploration strategy is reset to 1. As shown in Fig. 13 the sum of Q matrix starts decreasing as new states are explored and in most of the cases they do not meet the new target that is computed by (10). Thus, an agent receives often negative reward. As the amount of iterations increases and an agent learns the new environment, the cumulative reward increases and finally converges again.

As can be seen, the integration of Q-learning in the mLTE-U scheme can be of great importance towards the provision of fair coexistence between LTE and Wi-Fi in unlicensed spectrum. Q-learning can render an mLTE-U network capable to operate autonomously by learning and adapting into a dynamic wireless environment.

D. COMPARISON OF THE PROPOSED Q-LEARNING WITH CONVENTIONAL SELECTION SCHEMES

In this section, we compare the coexistence of mLTE-U with Wi-Fi, when mLTE-U selects the optimal configuration parameters using Q-learning with the case that mLTE-U is configured using conventional selection schemes, such as random and round-robin selection. According to the random selection scheme, mLTE-U configures the TXOP and muting period by selecting random values. For this scheme, uniformly distributed random selection is used. When round-robin is used, mLTE-U selects consecutively all the different configurations of TXOP and muting period. Such conventional mechanisms require lower complexity than Q-learning, as Q-learning must first learn the environment in order to offer optimal configurations. For this comparison and similar to Section VIII-B.1, we consider a scenario, in which one mLTE-U network coexists with one Wi-Fi network.

Fig. 14 presents the histogram of mLTE-U and Wi-Fi throughput for all the examined selection mechanisms. Fig. 14 (a) and Fig. 14 (d) show the respective histogram of the mLTE-U and Wi-Fi throughput according to the Q-learning mechanism. Fig. 14 (b) and Fig. 14 (e) present respectively the histogram of mLTE-U and Wi-Fi throughput when random selection mechanism is used and Fig. 14 (c) and Fig. 14 (f) illustrate the histogram of the corresponding mLTE-U and Wi-Fi throughput when round-robin selection mechanism is used. For every scenario, the throughput is calculated for the same number of iterations (7000 iterations). In every figure, the x-axis holds the obtained throughput value in Mbps, classified into series of intervals. The y-axis holds the frequency of the throughput value, meaning in how many iterations the obtained throughput value was in a specific interval.

As can be observed, in both random and round-robin mechanisms, the obtained throughput of mLTE-U and Wi-Fi is spread over all the possible values of the throughput that can be achieved by each network. This is related to the nature of the selection schemes, as the random scheme chooses in every interval a random pair of TXOP and muting period, while the round-robin scheme selects consecutively all the available combinations (serially one pair in each interval). Furthermore, in Fig. 14, it can be seen that the histograms of the random and the round-robin mechanisms are similar. This is related to the high number of iterations. In a long term basis and due to the uniformly distributed randomness, the random scheme selects every combination of TXOP and muting period for almost equal amount of times.

The supremacy of the proposed Q-learning scheme over the conventional schemes can be clearly seen in the graphs (a) and (d). As shown, using Q-learning the mLTE-U network learns the optimal configuration parameters that offer fair coexistence with the co-located Wi-Fi network. During the first iterations of Q-learning that correspond to the exploration phase the obtained throughput of mLTE-U and Wi-Fi varies, as very often random actions are chosen due to the high ϵ value. As the agent learns the environment and the value of ϵ decreases, the exploitation phase increases. As result, the agent chooses more and more often configuration values that approach the target value ($Thr_{target}^{\text{mLTE-U}}$) of the mLTE-U throughput. Hence, the dominant majority of the obtained mLTE-U throughput approaches its target value (72.64 *Mbps*), offering fair coexistence to Wi-Fi that achieves also half of its maximum throughput (15.4 *Mbps*).

IX. CONCLUSION

In our days and towards 5G, the number of heterogeneous networks increases rapidly. These networks consist of diverse wireless technologies with different requirements. The introduction of LTE-U has pushed the wireless community to find solutions that can enable fair coexistence of LTE with other well-established technologies in unlicensed spectrum. Towards a global solution that respects the regional requirements worldwide, 3GPP announced the LTE LAA standard according to which, LTE can operate in unlicensed spectrum through a secondary cell and by performing a CCA procedure before a transmission.

However, the ratio of transmission opportunities between LTE LAA and Wi-Fi is not balanced, especially in the case that Wi-Fi does not use frame aggregation. In order to enhance the fairness of LTE-U, an adaptable scheme named mLTE-U has been proposed. According to mLTE-U, LTE can transmit in unlicensed spectrum using an adaptable TXOP after a successful CCA. A TXOP is followed by an adaptable muting period. This muting period can be exploited by other co-located networks in order to gain access to the wireless medium.

In this article, we analytically study the mLTE-U scheme. The system model of mLTE-U, when it coexists with Wi-Fi is analyzed. Additionally, we introduce a Q-learning technique that can be used by an mLTE-U network to learn the wireless environment and autonomously select the TXOP and muting period configurations that can provide fair coexistence with other co-located technologies. Simulation results show how Q-learning can assist mLTE-U to find optimal configurations and be adapted to changes of the wireless environment providing the desired fair coexistence. Furthermore, the proposed scheme is compared with conventional selection schemes, revealing its supercity in providing fair coexistence with Wi-Fi.

In the near future, this work can be extended by exploiting deep Q-learning using neural networks, towards optimal selection of the mLTE-U parameters that can offer fair coexistence between mLTE-U and other co-located wireless technologies.

REFERENCES

[1] A. Bleicher, "A surge in small cells [2013 tech to watch]," *IEEE Spectr.*, vol. 50, no. 1, pp. 38–39, Jan. 2013, doi: 10.1109/MSPEC.2013.6395306.

[2] C. Cano and D. J. Leith, "Coexistence of WiFi and LTE in unlicensed bands: A proportional fair allocation scheme," in *Proc. IEEE Int. Conf. Commun. Workshop (ICCW)*, Jun. 2015, pp. 2288–2293, doi: 10.1109/ICCW.2015.7247522.

[3] "Making the Best Use of Unlicensed Spectrum for 1000x," Qualcomm, San Diego, CA, USA, White Paper, Sep. 2015. [Online]. Available: https://www.qualcomm.com/media/documents/files/whitepaper-making-the-best-use-of-unlicensed-spectrum.pdf

[4] *Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Layer Procedures Release 13*, document TS 36.213 3GPP, New Orleans, USA, Sep. 2016.

[5] *Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Layer Procedures Release 14*, document TS 36.213, 3GPP, Sapporo, Japan, Sep. 2017.

[6] *MulteFire Alliance*. Accessed: Feb. 2, 2018. [Online]. Available: http://www.multefire.org/

[7] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, Standard 802.11, New York, NY, USA, Mar. 2012.

[8] S. A. Rajab, W. Balid, M. O. Al Kalaa, and H. H. Refai, "Energy detection and machine learning for the identification of wireless MAC technologies," in *Proc. Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, Aug. 2015, pp. 1440–1446, doi: 10.1109/IWCMC.2015.7289294.

[9] V. Maglogiannis, D. Naudts, A. Shahid, and I. Moerman, "An adaptive LTE listen-before-talk scheme towards a fair coexistence with Wi-Fi in unlicensed spectrum," in *Telecommunication Systems*. New York, NY, USA: Springer, Jan. 2018, pp. 1–21, doi: 10.1007/s11235-017-0418-9.

[10] C. J. C. H. Watkins and P. Dayan, "Technical note: Q-learning," *Mach. Learn.*, vol. 8, pp. 279–292, May 1992.

[11] V. Maglogiannis, D. Naudts, P. Willemen, and I. Moerman, "Impact of LTE operating in unlicensed spectrum on Wi-Fi using real equipment," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Washington, DC, USA, Dec. 2016, pp. 1–6, doi: 10.1109/GLOCOM.2016.7841884.

[12] *iLab.t IMEC LTE Testbed*. Accessed: Feb. 2, 2018. [Online]. Available: http://doc.ilabt.iminds.be/ilabt-documentation/

[13] Y. Jian, C.-F. Shih, B. Krishnaswamy, and R. Sivakumar, "Coexistence of Wi-Fi and LAA-LTE: Experimental evaluation, analysis and insights," in *Proc. IEEE Int. Conf. Commun. Workshop (ICCW)*, Jun 2015, pp. 2325–2331, doi: 10.1109/ICCW.2015.7247528.

[14] R. Ratasuk *et al.*, "License-exempt LTE deployment in heterogeneous network," in *Proc. IEEE Int. Symp. Wireless Commun. Syst. (ISWCS)*, Paris, France, Aug. 2012, pp. 246–250, doi: 10.1109/ISWCS.2012.6328367.

[15] T. Nihtila *et al.*, "System performance of LTE and IEEE 802.11 coexisting on a shared frequency band," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Shanghai, China, Apr. 2013, pp. 1038–1043, doi: 10.1109/WCNC.2013.6554707.

[16] E. Almeida *et al.*, "Enabling LTE/WiFi coexistence by LTE blank subframe allocation," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2013, pp. 5083–5088, doi: 10.1109/ICC.2013.6655388.

[17] V. Maglogiannis, D. Naudts, A. Shahid, S. Giannoulis, E. Laermans, and I. Moerman, "Cooperation techniques between LTE in unlicensed spectrum and Wi-Fi towards fair spectral efficiency," *Sensors*, vol. 31, no. 9, p. 1994, Aug. 2017, doi: 10.3390/s17091994.

[18] C. K. Kim, C. S. Yang, and C. G. Kang, "Adaptive listen-before-talk (LBT) scheme for LTE and Wi-Fi systems coexisting in unlicensed band," in *Proc. 13th IEEE Annu. Consum. Commun. Netw. Conf. (CCNC)*, Jan. 2016, pp. 1–6, doi: 10.1109/CCNC.2016.7444845.

[19] A. Bhorkar, C. Ibars, A. Papathanassiou, and P. Zong, "Medium access design for LTE in unlicensed band," in *Proc. IEEE Wireless Commun. Netw. Conf. Workshops (WCNCW)*, Jun. 2015, pp. 369–373, doi: 10.1109/WCNCW.2015.7122583.

[20] F. Hao, C. Yongyu, H. Li, J. Zhang, and W. Quan, "Contention window size adaptation algorithm for LAA-LTE in unlicensed band," in *Proc. Int. Symp. Wireless Commun. Syst. (ISWCS)*, Poznan, Poland, Sep. 2016, pp. 476–480, doi: 10.1109/ISWCS.2016.7600951.

[21] V. Mushunuri, B. Panigrahi, H. K. Rath, and A. Simha, "Fair and efficient listen before talk (LBT) technique for LTE licensed assisted access (LAA) networks," in *Proc. IEEE 31st Int. Conf. Adv. Inf. Netw. Appl. (AINA)*, Taipei, Taiwan, Mar. 2017, pp. 39–45, doi: 10.1109/AINA.2017.135.

[22] B. Chen, J. Chen, Y. Gao, and J. Zhang, "Coexistence of LTE-LAA and Wi-Fi on 5 GHz with corresponding deployment scenarios: A survey," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 1, pp. 7–32, 1st Quart., 2017, doi: 10.1109/COMST.2016.2593666.

[23] Y. Li, T. Zhou, Y. Yang, H. Hu, and M. Hamalainen, "Fair downlink traffic management for hybrid LAA-LTE/Wi-Fi networks," *IEEE Access*, vol. 5, pp. 7031–7041, Dec. 2016, doi: 10.1109/ACCESS.2016.2642121.

[24] N. Rupasinghe and I. Güvenç, "Reinforcement learning for licensed-assisted access of LTE in the unlicensed spectrum," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, New Orleans, LA, USA, May 2015, pp. 1279–1284, doi: 10.1109/WCNC.2015.7127653.

[25] A. Shahid, S. Aslam, H. S. Kim, and K.-G. Lee, "A docitive Q-learning approach towards joint resource allocation and power control in self-organised femtocell networks," *Trans. Emerg. Telecommun. Technol.*, vol. 26, no. 2, pp. 216–230, Feb. 2014, doi: 10.1002/ett.2800.

[26] O. Sallent, J. Pérez-Romero, R. Ferrús, and R. Agustí, "Learning-based coexistence for LTE operation in unlicensed bands," in *Proc. IEEE Int. Conf. Commun. Workshop (ICCW)*, London, U.K., Jun. 2015, pp. 2307–2313, doi: 10.1109/ICCW.2015.7247525.

[27] A. Galanopoulos, F. Foukalas, and T. A. Tsiftsis, , "Efficient coexistence of LTE with WiFi in the licensed and unlicensed spectrum aggregation," *IEEE Trans. Cognit. Commun. Netw.*, vol. 2, no. 2, pp. 129–140, Jul. 2016, doi: 10.1109/TCCN.2016.2594780.

[28] O. Sharon and Y. Alpert, "MAC level Throughput comparison: 802.11ac vs. 802.11n," *Phys. Commun.*, vol. 12, pp. 33–49, Sep. 2014, doi: 10.1016/j.phycom.2014.01.007.

[29] A. Saif, M. Othman, S. Subramaniam, and N. A. A. Hamid, "Frame aggregation in wireless networks: Techniques and issues," *IETE Tech. Rev.*, vol. 28, no. 4, pp. 336–350, Jul. 2011, doi: 10.4103/0256-4602.83554.

[30] W. Liu, M. Kulin, T. Kazaz, A. Shahid, and I. D. P. E. Moerman, "Wireless technology recognition based on RSSI distribution at sub-Nyquist sampling rate for constrained devices," *Sensors*, vol. 17, no. 9, p. 2081, Sep. 2017, doi: 10.3390/s17092081.

[31] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: MIT Press, Feb. 1998.

VASILIS MAGLOGIANNIS received the M.Eng. degree in computer engineering and the M.Sc. degree in science and technology of computers, telecommunications and networks from the University of Thessaly, Greece, in 2012 and 2014, respectively. He is currently pursuing the Ph.D. degree with the Faculty of Engineering and Architecture, University of Ghent, Belgium. In 2014, he joined the IDLab, Department of Information Technology, University of Ghent–imec, Belgium. He has experience in working on national, FP7, and H2020 European Projects. His main research interests include mobile and wireless networks, 5G, LTE, 802.11, 802.15.4, and wireless sensor networks.

DRIES NAUDTS received the M.Sc. degree in computer science from the University of Ghent, Belgium, in 2001. He started as a Software Engineer with Kresoft and Nisus. In 2005, he joined the IDLab Research Group, Department of Information Technology, University of Ghent–imec, where he is currently a Senior Researcher. He is involved in several national, international Research and Development, FP7, and H2020 European projects. His research interests include mobile networks, 5G, LTE, 802.11, and C-ITS.

ADNAN SHAHID (SM'17) received the B.Eng. and M.Eng. degrees in computer engineering with wireless communication specialization from the University of Engineering and Technology at Taxila, Taxila, Pakistan, in 2006 and 2010, respectively, and the Ph.D. degree in information and communication engineering from Sejong University, South Korea, in 2015. He is currently a Post-Doctoral/Senior Researcher with the IDLab, Department of Information Technology, University of Ghent–imec, Belgium, where he is involved in several European research projects, including eWINE, WiSHFUL, and so on, and other national projects, including SAMURAI, IDEAL-IOT, MUSCLE-IOT, and so on. He is also technically coordinating the eWINE Project. From 2015 to 2016, he was with the Department of Computer Engineering, Taif University, Saudi Arabia. In 2015, he was a Post-Doctoral Researcher with Yonsei University, South Korea, where he was also a recipient of the prestigious BK21 plus post-doctoral and research professor positions. He is actively involved in various research activities. He has over 30 research papers in well-known venues at journals, conferences, patents, and so on. His research interests include the next-generation wireless communication and networks with prime focus on resource management, interference management, cross-layer optimization, self-organizing networks, small-cell networks, device-to-device communications, machine-to-machine communications, Sug-GHz technologies, 5G wireless communications, machine learning, and so on. He is serving as an Associate Editor for the IEEE ACCESS Journal and the *Journal of Networks and Computer Application* (Elsevier).

INGRID MOERMAN received the degree in electrical engineering and the Ph.D. degree from Ghent University in 1987 and 1992, respectively. She became a part-time Professor with Ghent University in 2000. She is currently a Staff Member with the IDLab, a core research group of imec with research activities embedded in Ghent University and the University of Antwerp. She is also coordinating the research activities on mobile and wireless networking, and she is also leading a research team of about 30 members with the IDLab–Ghent University. She has a longstanding experience in running and coordinating national and EU research funded projects. At the European level, she is, in particular, very active in the Future Networks research area, where she has coordinated and is coordinating several FP7/H2020 projects (CREW, WiSHFUL, eWINE, and ORCA) and participating in other projects (Fed4FIRE, FORGE, FLEX, and Flex5Gware). She has authored or co-authored of over 700 publications in international journals or conference proceedings. Her main research interests include Internet of Things, low power wide area networks, high-density wireless access networks, collaborative and cooperative networks, intelligent cognitive radio networks, real-time software-defined radio, flexible hardware/software architectures for radio/network control and management, and experimentally supported research.