LTE on License-Exempt Spectrum

Jingjing Zhang, Mao Wang, Min Hua, Tingting Xia, Wenjie Yang, and Xiaohu You, *Fellow, IEEE*

***Abstract*—The radio frequency spectrum is classified as licensed and license-exempt/unlicensed spectrum. A traditional cellular communications system (e.g., LTE) operates unexceptionally on licensed spectrum. This paper explains the concept of cellular communications on both licensed and license-exempt carriers under a unified cellular architecture. It addresses the key challenges and describes a baseline communication framework that enables such operation, including carrier sensing based on listen-before-talk, discontinuous transmissions with limited channel occupation time, synchronization between licensed and license-exempt carriers, and coexistence with other incumbent systems like WiFi. These concept and techniques are further exemplified by a practical system, the LTE license-assisted access, including the downlink featured in Release 13 LTE and the most recent addition of uplink in Release 14 LTE.**

***Index Terms*—Cellular communications, LTE, control channel, traffic channel, radio spectrum, licensed spectrum, license-exempt/unlicensed spectrum.**

I. Introduction

THE STRONG trend of media consumption moving toward the mobile devices continues to pressure wireless connectivity operators to deliver ever-growing volumes of mobile data to meet the ever-increasing demand for faster mobile data services (e.g., the fiber-like user experience [1]). Challenges of this magnitude clearly require not only new wireless technologies but also *new spectrum*. Today's cellular networks, like the LTE network [the Third-Generation Partnership Project (3GPP) in the forms of its Long-Term Evolution project], are already operating at a very high spectral efficiency, leaving little margin for further practical and cost-effective improvements. Extremely-densely deployed small cells towards so called hyper-dense heterogeneous networks are one of the main solutions to the challenge by increasing the area spectral efficiency or frequency reuse [2]. Currently, cellular small cell networks are unexceptionally operating on the scarce licensed bands. Thus, the network capacity is ultimately upper-bounded by the availability of these licensed bands. More spectrums are apparently the ultimate solution, and are needed more than ever. Therefore, we need to look into spectrum bands beyond the limited licensed spectrum, such as the spectrum-rich 5 GHz license-exempt band for potential data rate boost as graphically depicted in Figure 1 at the top of the next page [3].

License-exempt or simply "unlicensed" spectrum is increasingly considered by cellular operators as a complementary tool to off-load data traffic from the congested licensed cellular bands and boost overall system throughput. The approaches to cellular communications on unlicensed spectrum have been extensively studied recently. The fundamental framework of cellular communications on unlicensed spectrum and a conceptual downlink system have been introduced in [4]. Two practical cellular technologies for communications on unlicensed spectrum, the LTE-U and the LTE license-assisted access (LAA), have been respectively investigated by LTE-U Forum and 3GPP LTE. A high-level overview of LTE-LAA as specified in Releases 13 and 14 LTE is provided in [5] and [6]. Potential benefits and the issues of fair coexistence of LTE-U/LAA with the incumbent unlicensed band users are discussed in [7] and [8] from a high-level perspective. Various coexistence algorithms and analyses are presented in [9]–[11].

Related studies are also available in the literature. A survey of the state-of-the-art spectrum occupancy models for cognitive radio designs is given in [12] and [13] and an overview of the spectrum prediction techniques are studied in [14]. A survey of spectrum measurement techniques and associated interference maps as well as means to improve measurement accuracy is provided in [15]. Tehrani *et al.* [16] provides a survey of the licensed spectrum sharing regimes to fully exploit the licensed spectrum. The potential deployment scenarios, benefits and challenges are also elaborated. The concept of device-to-device (D2D) links as a new cell tier to wireless networks in licensed bands has been explored over the past several years. A detailed review of the cellular D2D in a licensed band is given in [17], which includes resource management, interference management, power consumption in different topology scenarios (such as broadcast and relay). The coexistence of D2D communication in small cells is also studied and the analytical results of network assisted-D2D on unlicensed spectrum are given in [18], in which devices are continually associated with the cellular network and use this connectivity to help manage their D2D connections in the unlicensed bands. D2D on unlicensed band with the assistance from cellular network is further investigated to show the potential improvement to the overall network capacity in [19]. A complete overview and the engineering details of cognitive radios, the heterogeneous network model, and power and cost challenges in the context of future machine-to-machine (M2M) cellular networks are shown in [20].

Manuscript received February 13, 2017; revised June 14, 2017 and September 10, 2017; accepted October 24, 2017. Date of publication November 8, 2017; date of current version February 26, 2018. This work was supported in part by the National Natural Science Foundation of China under Grant 61771128, and in part by the Open Research Fund of the National Communications Research Laboratory, Southeast University under Grant 2017B01. *(Corresponding author: Mao Wang.)*

J. Zhang, M. Hua, T. Xia, and W. Yang are with the School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China.

M. Wang and X. You are with the School of Information Science and Engineering, Southeast University, Nanjing, China (e-mail: wangmao@seu.edu.cn).

Digital Object Identifier 10.1109/COMST.2017.2771485

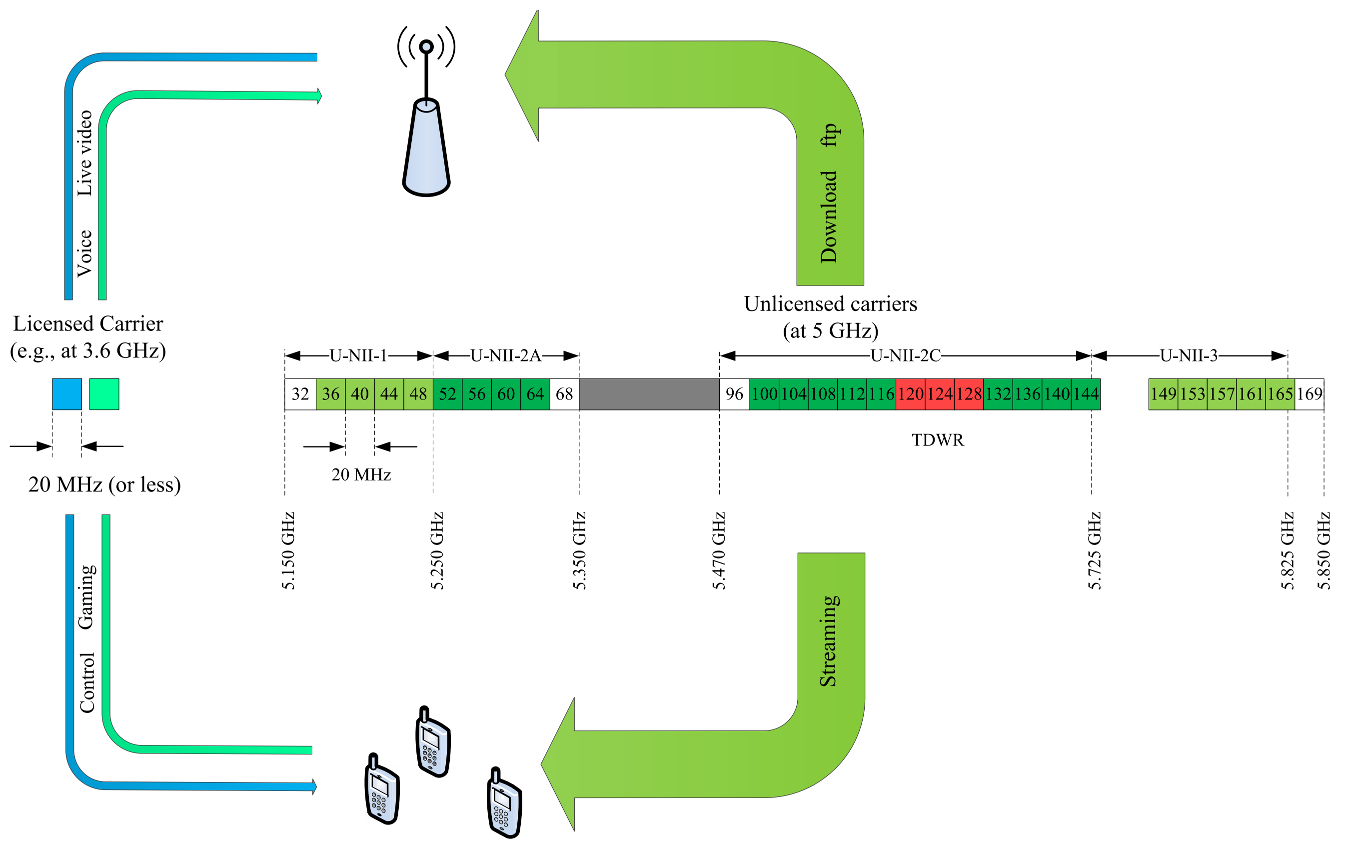


Fig. 1. Channelization of 5 GHz unlicensed spectrum. Also, a unified cellular deployment scenario is depicted to exploit the full benefits of the centrally coordinated and managed cellular architecture, and unlicensed bands. QoS-crucial data services (e.g., voice, live video, and gaming) are delivered through licensed carrier, whereas latency-insensitive data (e.g., ftp download) are delivered through unlicensed carriers (currently on unlicensed national information infrastructure (U-NII)-1 and U-NII-3 only, where no procedure for the dynamic frequency selection spectrum-sharing with radar systems is required). The control signals/messages are through LTE cellular infrastructure. In essence, this deployment model leverages the large number of small cells to work as a unified cellular network to efficiently exploit both licensed and unlicensed spectrum bands. The green-colored channels are unlicensed bands with minimum bandwidth of 20 MHz, of which the channels in dark green require the dynamic frequency selection spectrum-sharing mechanism to ensure coexistence with radar systems. TDWR channels (Channels 120,124, and 128) are Terminal Doppler Weather Radar channels (Federal Aviation Administration) that are not allowed to be used for other purposes.

The purpose of this paper is to give a comprehensive analysis on the physical layer design principles of cellular communications on unlicensed spectrum with applications in the LTE system, i.e., LTE-LAA including the most recent addition of the unlicensed uplink in Release 14 LTE [6], [21], [22].

This paper is organized as follows: Section II reviews the regulatory requirements on unlicensed spectrum, including the distinctions between communications on licensed and on unlicensed spectrum, particularly, the channel sensing based on the listen-before-talk (LBT) rule (frame-based or load-based) imposed in certain regions for better coexistence. Section III addresses the benefits and challenges and Section IV is devoted to the analysis of the key factors associated with implementation of cellular communication structure on unlicensed spectrum, such as frame-based or load-based LBT, synchronous or asynchronous LBT, and the use of a reservation signal. A baseline cellular framework that facilitates the LBT is then derived. With Section IV as the foundation, Section V describes two practical systems, LTE-U and LTE-LAA. The former provides a simple, yet effective solution for an existing LTE system to operate on the unlicensed spectrum in regions where LBT is not enforced; and the latter is a unified approach to ultimately realizing the LBT functions on the LTE architecture. The detailed design aspects and rationale behind such practical systems are described with a focus on the latest industrial activities in extending LTE into the unlicensed spectrum, particularly the 3GPP LAA in Release 13 and Release 14 LTE. Final notes on future research directions including the hidden node problem, and machine-type communications (MTC) and D2D on unlicensed spectrum are given in Section VI. Finally, Section VII concludes the paper. We summarize the definitions of the acronyms that will be frequently used in this paper in TABLE I for ease of reference.

II. Unlicensed Spectrum and Regulatory Requirements

The radio frequency (RF) spectrum that has been taken into utilization for communication services ranges from 10 kHz to 80 GHz and is extending to higher frequencies. The radio spectrum regulator such as the Federal Communications Commission (FCC) in the United States (Region 2) uses several mechanisms to make spectrum available for communication services through licensed and license-exempt (or unlicensed) spectrum.

TABLE I
ABBREVIATION

CCA	clear channel assessment
CP	cyclic prefix
CR	cognitive radio
CRS	cell-specific reference signal
CSI	channel state information
CSMA/CA	carrier-sense multiple access with collision avoidance
CTS	clear-to-send
D2D	device-to-device
DCF	distributed coordination function
DFS	dynamic frequency selection
DMRS	demodulation reference signal
DRS	discovery reference signal
DwPTS	downlink pilot time slot
EIRP	equivalent effective isotropically radiated power
eNodeB	evolved Node B (LTE base station)
EPDCCH	enhanced physical downlink control channel
FCC	Federal Communications Commission
FDM	frequency-division multiplexing
HARQ	hybrid automatic repeat request
ISM	industrial, scientific, and medical
ITU	International Telecommunications Union
LAA	license assisted access
LTE	(3rd Generation Partnership Project) Long Term Evolution
LTE-A	LTE-Advanced
LBT	listen-before-talk
M2M	machine-to-machine
MAC	medium access control
MCOT	maximum channel occupancy time
mIoT	massive Internet-of-Things
MTC	machine-type communication
NAV	network allocation vector
NB-IoT	narrow band Internet-of-Things
OFDM	orthogonal frequency division multiplexing
PAPR	peak to average power ratio
PDCCH	physical downlink control channel
PDSCH	physical downlink shared channel
PHICH	physical HARQ indicator channel
PSD	power spectral density
PSS	primary synchronization signal
PUCCH	physical uplink control channel
PUSCH	physical uplink shared channel
QoS	quality of service
RB	resource block pair
RF	radio frequency
RRC	radio resource control
RRM	radio resource management
RSRP	reference signal received power
RSRQ	reference signal receiving quality
RTS	ready-to-send
SC-FDM	single carrier frequency division multiplex
SSS	secondary synchronization signal
TDM	time-division multiplexing
TDWR	Terminal Doppler Weather Radar
TTI	transmission time interval
UE	user equipment
U-NII	unlicensed national information infrastructure

Some licensed frequencies are designated for commercial use while others are allocated to public safety agencies. The assignment of a frequency band to a user and the issue of the associated license to the user give the user the authority for transmitting at that frequency (the assigned frequency) and the bandwidth of emission (the assigned frequency band) for stated purposes using stated emission parameters [23]. Licensed spectrum allows for exclusive use of particular frequencies in particular geographic locations, meaning when someone is granted the right by, e.g., FCC to communicate at certain frequencies and in certain locations, everyone else is prohibited from using that frequency in that location.

Whereas in spectrum that is designated as license-exempt or unlicensed, users can operate without a regulator license but must comply with the regulatory constraints. For instance, the regulator limits the transmit power and the equivalent effective isotropically radiated power (EIRP) in order to keep the interference to other co-channel systems acceptably low. Nevertheless, users on the unlicensed bands do not have exclusive use of the spectrum and are subject to interference. Apparently, given the power limits placed on transmitters, unlicensed spectrum would be most applicable to small cells.

The unlicensed frequency bands as listed in Table II are originally allocated for industrial, scientific, and medical (or ISM bands for short) applications. They were initially established at International Telecommunications Conference of International Telecommunications Union (ITU) in 1947. It is interesting to note that these bands originally were not intended to be used for wireless communications. In fact, just the opposite is true [24].

Indeed, apart from the radio communication services whose functioning depends on the radiation of the RF energy and its reception at a distance, there are other RF use cases that do not involve deliberate radiation outside the frequency limits of the equipment in which it is operated. These include ISM equipment, from which a considerable amount of RF energy may be radiated *unintentionally*, capable of causing severe interference to radio communications nearby. These ISM applications include everything from industrial heating equipment, medical diathermy machines to home-use microwave ovens. A number of frequency bands have been designated internationally for use by ISM equipment. This equipment is required to operate only in specific frequency bands in various parts of the radio spectrum, whose manufactures are required to control the radiation frequency within these bands without interfering radio communications operating in anywhere other than these bands. They collectively referred to as the ISM bands. Table II lists the principal ISM bands. Other bands are also available for ISM use but subject to the special authorization of the regulator. These bands are 6.765-6.795 GHz, and 61.0-61.5 GHz, for example. In addition, the band 433.05 – 434.79 MHz is available for ISM applications in some European countries and some countries in Region 1. ITU-R reviews the principal ISM applications and lists some typical levels of leaked radiations [24]. There are no agreed constraints on radiation from ISM apparatus.

Apparently, communications are probably the last things one would expect in these bands. Nevertheless, despite the original intention, radio communications in the ISM bands are possible as long as the communication systems are designed to tolerate the ISM interference as well as the interference potentially from other communication systems operating in the same band. Ever since the advent of mobile devices, more and more short range, low-power, low-cost wireless communication systems like cordless phones, WiFi [25], Bluetooth [26], and ZigBee [27] have found their homes in some of these unlicensed bands like 902 – 928 MHz, 2.40 – 2.4835 GHz, and 5.725 – 5.875 GHz.

TABLE II
ISM BANDS

Band name	ITU band	Frequency allocation
High Frequency (HF)	7	6.765–6.795 MHz
High Frequency (HF)	7	13.553–13.567 MHz
High Frequency (HF)	7	26.957–27.283 MHz
Very High Frequency (VHF)	8	40.66–40.70 MHz
Ultra-High Frequency (UHF)	9	902 – 928 MHz (ITU Region 2)
Ultra-High Frequency (UHF)	9	2.40 – 2.4835 GHz
Super High Frequency (SHF)	10	5.725 – 5.875 GHz

The growing wireless applications in the ISM bands gave an immense impetus to the wireless industry to increase the amount of spectrum available for unlicensed use. In 1997, FCC made available 300 MHz of spectrum at 5.15–5.25 GHz (U-NII-1), 5.25–5.35GHz (U-NII-2A), including 5.725–5.825GHz (referred to as U-NII-3), for use by a new category of unlicensed equipment [28]. U-NII-3 is partially overlapped with the ISM band (5.725 – 5.875 GHz), and hence is sometimes referred to as U-NII/ISM. In 2003, FCC made available additional 255 MHz worth of spectrum from 5.47 to 5.725 MHz (U-NII-2C). This aligns the U-NII frequency band in the United States with other parts of the world, thereby allowing the same product to be used in most parts of the world.

Note that the frequency bands 5.250–5.350 and 5.470–5.725 GHz in U-NII-2 are used by radar systems worldwide and the use of these bands requires the dynamic frequency selection (DFS) technique to support cognition over these bands [29]. As shown in Figure 1, the channels in dark green require the DFS spectrum-sharing mechanism to ensure coexistence with radar systems. DFS allows a U-NII-2 user, like a WiFi (particularly IEEE 802.11h) system, to detect and avoid co-channel operation with radar systems by automatically selects a frequency that does not interfere with certain radar systems while operating in the U-NII-2 bands. DFS detects a rising edge of in-band power, checks the signal bandwidth and frequency for matches against radar pulse characteristics, and waits for a falling edge of in-band power. The detected pulse is then further analyzed against a set of parameters, such as pulse width, timing, repetition interval, amplitude, modulation type (if any), total number of pulses, etc., for various radar types. A user verifies a channel is free of radar before using it, monitors for radar, and vacates the channel if a radar signal is detected.

In certain regions, such as the European Union and Japan (Region 1), the LBT rule is enforced to reduce the interference risk to others as well as themselves for better coexistence among different wireless systems operating on the same unlicensed band. The LBT medium access rule prevents a transmitter from continuous transmission and monopolizing the resource. Rather, it requires the transmitter to check for other occupants such that it does not impact other occupants.

In particular, the LBT medium access rule according to ETSI EN 301893 [29] requires that a transmitter waits for its turn if there is evidence that another transmitter is using the channel. A process called *clear channel assessment* (CCA) is used to determine if the channel is available for transmission. CCA checks the received energy for channel activities before transmitting. The minimum CCA observation time is 20μs. The basic assumption under CCA is that a packet being transmitted carries a signal in which the received signal energy is high enough to exceed a specified level. If the energy is detected exceeding the CCA threshold (i.e., CCA fails), the channel is assumed to be in use. Otherwise, the transmitter can transmit for duration equal to the *channel occupancy time* and at a bandwidth at least, e.g., 80% of the total bandwidth at 5 GHz, which imposes limits on the maximum and minimum duration, and minimum frequency band occupancy of a transmission burst.

III. BENEFITS AND CHALLENGES

In this section, we put forward the potential benefits and challenges associated with the implementation of cellular communications on the unlicensed spectrum.

A. Benefits

It is highly desirable to have a cost-effective solution for integrating several existing network resources (licensed or unlicensed, which are allocated for different services) into a single mobile converged network. Currently, existing wireless wide area networks like the cellular network (e.g., LTE operating on the licensed frequency bands) and the wireless local area network (WLAN) (e.g., WiFi operating on license-exempt or unlicensed bands) operate independently on the licensed and unlicensed bands. It is natural and beneficial to aggregate licensed carrier together with unlicensed carriers to boost throughput while maintaining support for seamless mobility and high reliability.

Cellular systems and WLAN systems employ fundamentally distinctive architectures to cope with different channel properties (due to the different regulatory rules as will be discussed in the next section), and to achieve different goals [4]. Cellular systems operating on licensed spectrum are characterized by high spectral efficiency, reliable and predicable data service performance, and robust mobility, whereas most WLAN systems on unlicensed bands are typically cost-effective and flexible in deployment but are often spectrally inefficient and lack of quality of service (QoS) control. A simple combination of LTE and WiFi to create a carrier WiFi hybrid system is clearly "sub-optimal" in terms of leveraging the potential multiplexing gains and overall resource usage efficiency. Since these are the two very different technologies and systems, interworking is more complex for both the network and the mobile devices. An intriguing option could just be to extend LTE into unlicensed spectrum which aggregates both licensed and unlicensed bands to provide an extension of a larger LTE network, allowing for seamless flow of data between licensed and unlicensed spectrum with the same technology through a single core network that employs the same authentication, operations and management systems, and the same acquisition, access, registration, paging and mobility procedures. This means reduced cost, lower overhead, strengthened system performance, and higher overall network capacity. Undoubtedly, the unification of network management, streamlining authentication, handover, tracking, and resource allocation is a strong motivating factor for unlicensed LTE rollout.

In the exemplary illustration of Figure 1, data streams are aggregated and carried on both licensed band (e.g., 20 MHz or less using LTE air interface) and unlicensed bands (e.g., ∼500 MHz spectrum at 5 GHz split into multiple 20 MHz bands using LTE-U/LAA air interface), wherein system control signaling and QoS-crucial data services (e.g., voice, live video, and gaming) are delivered through LTE licensed carrier, whereas mobile traffic for latency-insensitive applications like file download and video streaming are delivered *opportunistically* through LAA on unlicensed bands to meet capacity demand. This enables the LTE network to leverage the large number of LAA cells to work as a unified network to efficiently exploit both licensed and unlicensed spectrum bands.

B. Challenges

Despite the advantages of unlicensed LTE, there are also serious technical challenges implied by this concept. In a traditional cellular system operating on licensed spectrum, the network has the right of exclusive use of the spectrum. Multiple technologies can coexist only in a primary–secondary dynamic spectrum access mode. Therefore, for the primary user (e.g., the cellular network), the utilization of radio resources is guaranteed and there is no competition. This allows transmission/reception to be organized into a highly efficient frame structure that is continuous and follows a deterministic timing. With this contiguous frame structure, a network can continuously utilize the resources and efficiently manage them without sharing with other systems. The only interference is from its friendly/cooperative neighboring cells belonging to the same network, which is mitigated through network planning [30] and/or coordination among cells [31], [32]. On the contrary, in an unlicensed band there is no guaranteed use of resources. Resources are used in an autonomous and competitive manner and, consequently, stringent QoS requirements are difficult to be ensured because of uncontrolled interference. The challenge is how to apply the cellular technologies to unlicensed spectrum; how to offer reliable mobile services in these unreliable bands using the LTE technology and network.

Although it is true that unlicensed access of spectrum offers fair competition opportunities for multiple technologies, it is also important to assure the fair coexistence of new systems like LTE with incumbent systems like WiFi. The main coexistence issue arises from the difference in the design of the MAC layer mechanisms that WiFi and LTE implement. The WiFi MAC is based on Distributed Coordination Function (DCF), which is essentially a CSMA/CA-based LBT mechanism [33], [34]. As aforementioned, LTE is centralized scheduling based, whose transmission follows a continuous frame structure without yielding to any other system. Such a behavior works fine in the licensed band however would completely starve WiFi systems in unlicensed band.

For regions with LBT enforced for better coexistence, the LTE transmission structure is clearly prohibitive. However, due to the random nature of LBT, the implementation of LBT in the centralized scheduling transmission framework is probably one of the toughest challenges that the cellular technology has ever faced. There seems to be no easy solution. And yet the minimum spectrum occupation regulation may require the change of its multiple access waveform. Together, they may require significant changes in both MAC and physical layers. They are the major task for systems like LTE-LAA, and ergo the focus of this paper. We elaborate on these issues that emerge when analyzing the design of the LTE-U/LAA in the ensuing sections.

Indeed, the focus until now has been on the alteration of the traditional cellular architecture to provide the LBT functionality. Nevertheless, other challenges, such as coexistence of multiple LTE-U/LAA operators, hidden nodes issues, massive MTC for massive Internet-of-things (mIoT), and the implementation of D2D communications on the unlicensed spectrum are also of great concern although they have not been addressed in the current releases of LTE, and received only limited attention to date. They certainly are topics of great importance and urgency for future research.

IV. Baseline LBT Framework for Cellular System

In this section, we first briefly review the transmission structure of a typical cellular system. We then analyze the realization options for a cellular system complying with the regulatory requirements on the unlicensed spectrum with focus on LBT and its realization on the cellular transmission structure along with the techniques for co-existence with other systems while maximally preserving the cellular frame structure. The analytical results from this section pave the way for more detailed discussions of the LAA technology discussed in Section V.

A. Cellular Transmission Structure

As previously pointed out, cellular communications (e.g., LTE) utilize a deterministic and contiguous frame structure to facilitate centralized scheduling of resources (i.e., the assigned frequency band) among users [or user equipment (UE) in LTE terminology] within the system. With this structure, the network that owns the spectrum continuously utilizes the resources and efficiently manages them without interruptions from other systems. The continuous transmission structure also allows for the contiguous transmission of reference signals, such as the cell-specific reference signal (CRS), providing continuous time and frequency synchronization for receivers (UEs) as shown in Figure 2.

The transmission timeline of the LTE cellular system is composed of radio frames, each of which is further divided into 10 contiguous 1-ms subframes – the minimum transmission time intervals (TTIs). A subframe is further divided into two slots, each made of seven orthogonal frequency division multiplexing (OFDM) symbols. A resource block is defined as 12 subcarriers over one slot. The minimum resource unit that can be scheduled is 12 subcarriers of one subframe (one TTI), i.e., a pair of resource blocks, henceforth simply referred to as "RB". The downlink scheduling

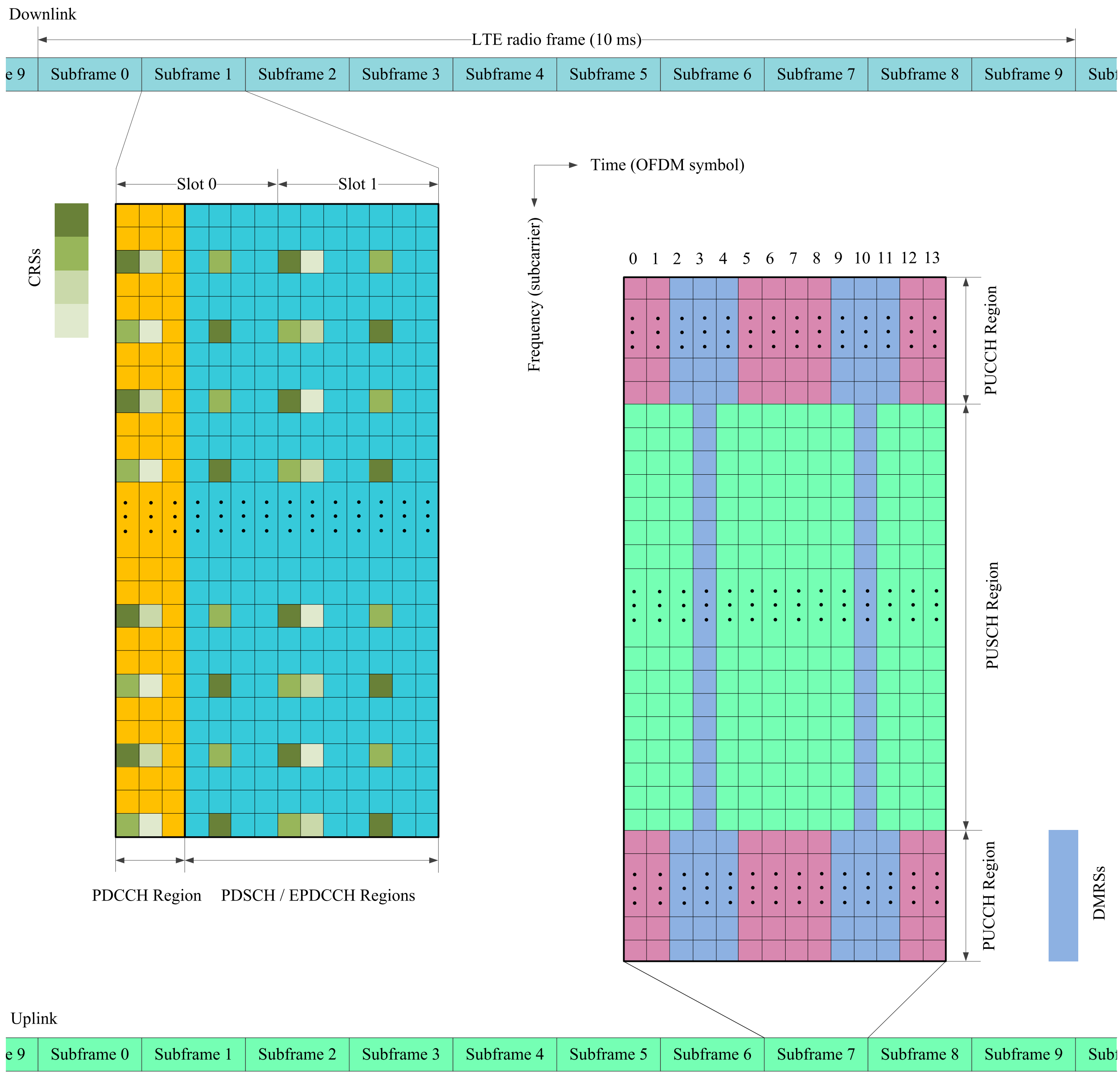


Fig. 2. Illustration of LTE frame structure, in which the transmission follows a continuous and deterministic time line that consists of a sequence of contiguous radio frames. A radio frame contains 10 subframes, each of which is composed of 14 OFDM symbols (i.e., two slots). The first up to three OFDM symbols of a downlink subframe are used for control channels like PDCCHs and PHICHs, and the remainder is primarily for downlink traffic (PDSCHs). The enhanced PDCCHs (EPDCCHs) can also be frequency-division-multiplexed with PDSCHs. The two end parts of an uplink subframe in frequency domain are used for PUCCH, while the rest is for uplink traffic (PUSCHs). DMRSs are demodulation reference signals.

information, i.e., the grants on the usage of the physical downlink shared channels (PDSCHs) of the current subframe is transmitted per subframe on the physical downlink control channel (PDCCH) or the enhanced PDCCH (EPDCCH). PDCCH resides on the first up to three OFDM symbols, multiplexed with its associated PDSCH of the current subframe in a time-division multiplexing (TDM) pattern within a subframe. EPDCCH adopts the same transmission resource unit (i.e., RB) as PDSCH, sharing the channel with the corresponding PDSCH in a frequency-division multiplexing (FDM) manner within a subframe [35].

The uplink scheduling information, i.e., the assignments on the usage of physical uplink shared channel (PUSCH), is also carried on the PDCCH or EPDCCH. The typical uplink scheduling delay is four subframes from the uplink assignment on PDCCH to the associated uplink data on PUSCH to accommodate the decoding of scheduling information on PDCCH, data preparation (e.g., rate matching and channel encoding), and transmission timing advance[1] [35].

[1]Timing advance on the uplink transmission is explained in detail in Section V-B *2)*.

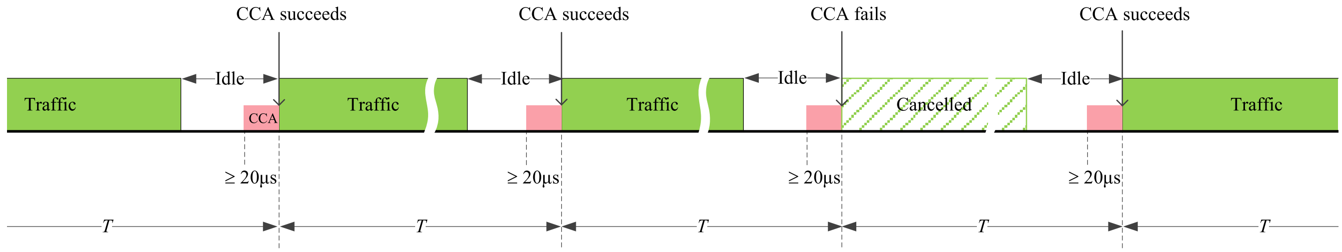


Fig. 3. Illustration of frame-based LBT, where transmission obeys a fixed timeline with a constant period T. CCA is performed before each transmission. An idle time of no less than 5% of channel occupancy time is required. The duration of CCA is fixed (e.g., 20μs) regardless of the outcome. A successful CCA warrants the transmission, whereas a failed CCA cancels the transmission.

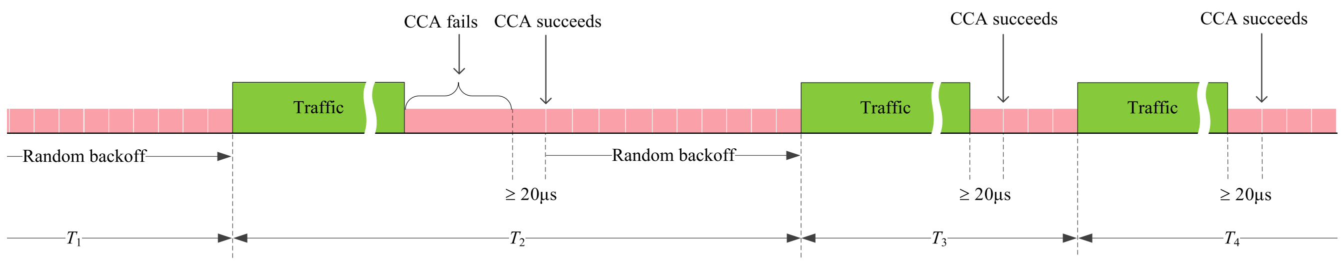


Fig. 4. Illustration of load-based LBT, where the CCA duration is random, resulting in a variable transmission timeline. The variability is due to the random backoff or CCA countdown. In addition, the potential interruption from other systems (not shown) may cause CCA to fail (shown), resulting in extensions in its duration.

The corresponding acknowledgement (ACK/NACK) to the downlink transmission on a PDSCH is fed back by the receiving UE via the physical uplink control channel (PUCCH), four subframes after the corresponding PDSCH. Asynchronous hybrid automatic repeat request (HARQ) is employed in the downlink for re-transmission (if necessary), in which an LTE base station or eNodeB provides explicit instructions to the UE regarding which RBs are used for the re-transmission via PDCCH collocated with the PDSCH used for the re-transmission in the same subframe. The UE needs to blindly detect the PDCCH in each subframe.

As for the uplink, ACK/NACK to the uplink transmission on a PUSCH is indicated to the UE via the downlink physical HARQ indicator channel (PHICH) transmitted four subframes after the associated PUSCH. Re-transmissions, if necessary, follow a synchronous HARQ timeline, which happen at a fixed-time interval, i.e., eight subframes apart.

Compared to the synchronous HARQ, asynchronous HARQ offers more flexibility in scheduling whereas requiring higher signaling overhead.

B. LBT Transmission Structures

There are two basic types of LBT transmission structures: the frame-based LBT and load-based LBT [29].

The frame-based LBT transmission structure is not demand-driven, but follows a deterministic timeline. Transmissions can only happen at specific times with minimum duration of, e.g., 1 ms and maximum duration of, e.g., 10 ms [29] as depicted in Figure 3. As a result, the CCA window has to be fixed as well, i.e., right before the transmission time. A transmitter relinquishes the channel and waits for the next CCA opportunity once the CCA fails. An idle period of more than 5% channel occupancy time is mandatory to leave an opportunity for other competitors to access the channel.

On the contrary, for load-based transmission as illustrated in Figure 4, the transmission time is not fixed. The CCA is performed continuously without abiding any frame boundaries until succeeds. At this time, a random backoff (or CCA countdown) timer is set off to perform an extended CCA to introduce randomness among competitors for collision avoidance. The random-backoff timer is decremented when a CCA slot succeeds; otherwise, the timer remains frozen. Transmission starts as soon as the timer expires.

1) Frame-Based LBT: A straightforward implementation of the frame-based LBT (see Figure 3) in the LTE frame structure is depicted in Figure 5 (a). The first subframe of a radio frame (Subframe 0) is designated for LBT, of which the first part is devoted to the idle period to satisfy the idle period requirement (e.g., 5%) by the frame-based LBT, and the second part (the last few OFDM symbols) are used for CCA. It is not difficult to see that this implementation suffers from perpetual collisions among multiple synchronous frame-based LBT systems that all clear the CCA.

This issue can be solved by introducing randomness among CCA performed by different systems. Specifically, as shown in Figure 5 (b), the first half of Subframe 0 is devoted to the idle period in which no transmission is permitted. Hence, the time duration of this period is 0.5 ms. The second half is reserved for implementing the CCA, which is further subdivided into 20 CCA slots for contention, i.e., for transmitters to compete for the usage of the following nine subframes.

The CCA opportunity (CCA slot) in the contention window, within which a cell performs CCA, is given by a random number generator with a given seed. In different contention windows, a cell obtains different CCA slot from the generator,

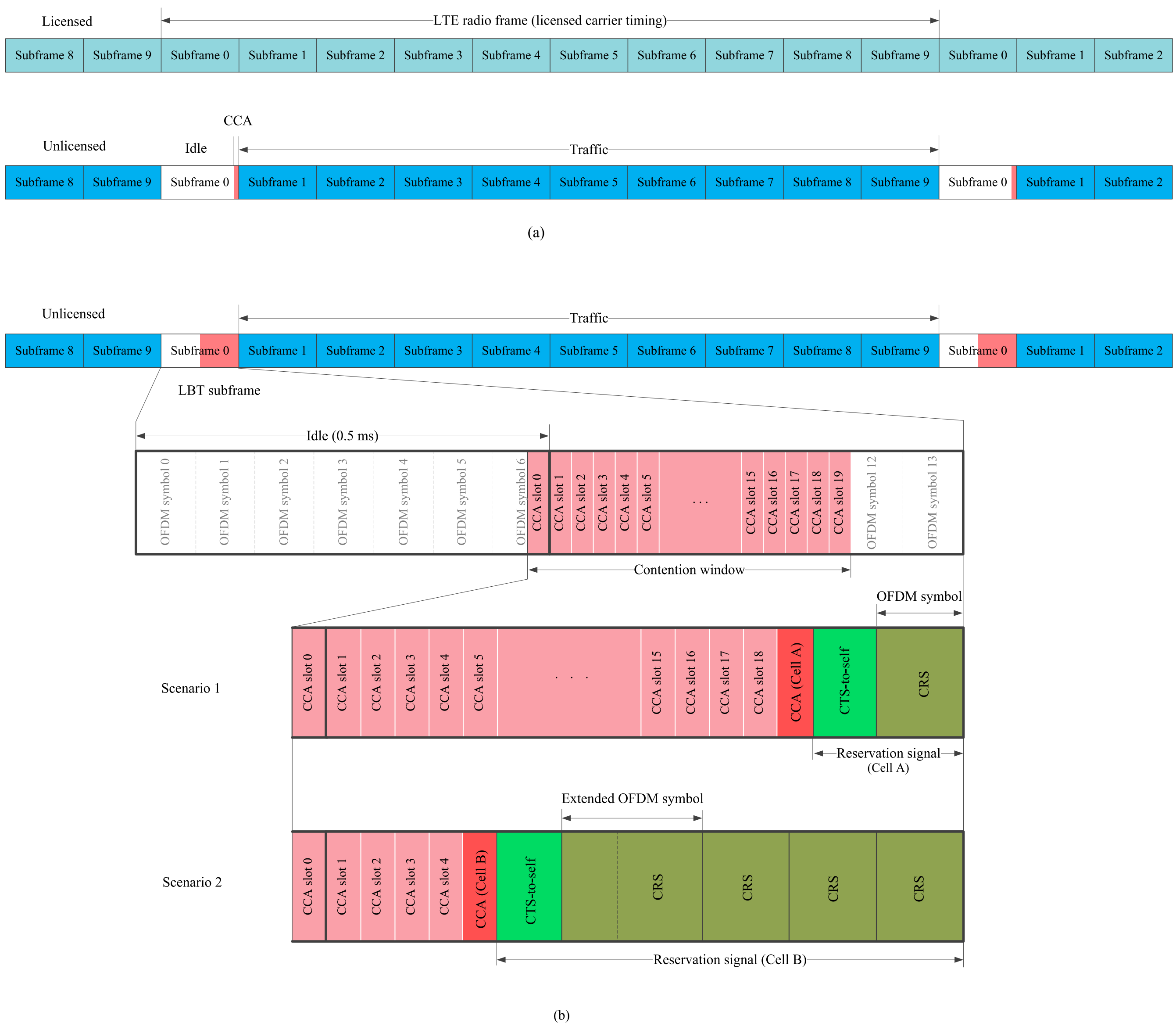


Fig. 5. (a) An implementation of the frame-based LBT under the LTE frame structure, where Subframe 0 is devoted to LBT, and Subframes 1 through 9 are reserved for traffic. (b) An augmented frame-based LBT. The first slot of Subframe 0 is devoted to the idle period in which no transmission is permitted to satisfy the 5% idle period requirement. Each CCA slot occupies $\sim 20\mu$s. In scenario 1, only one system is present in the operating channel. Cell A is given a CCA opportunity at CCA slot 19 (by a random number generator ranging from 0 to 19). Cell A transmits a reservation signal immediately to secure the channel as soon as it wins the CCA (the medium is clear since only one system is present). In scenario 2, two systems are present in the channel. Cell B is given a CCA opportunity at CCA slot 5 and succeeds (the medium is clear in slot 5) and grabs the channel by transmitting a reservation signal immediately, which silences Cell A (Cell A will fail the CCA in slot 19) till the next LBT subframe in the next radio frame when both cells will have new CCA slots to compete for the medium again.

ensuring fairness among different cells. The seed can be either cell-specific or network-specific. In the latter case, all cells belonging to the same network share the same CCA slot, and hence do not block each other. This allows full frequency reuse in the same network, as in a traditional LTE cellular network.

As the example shown in Figure 5 (b), Cell A is given CCA slot 19. Cell A thus listens on the channel for activities by performing CCA during the 19th CCA slot. In this case, since Cell A is the only system present, the measured channel energy is presumably below the specified CCA threshold. Cell A thus takes over the following nine subframes (Subframes 1-9), and secures it by immediately transmitting a "reservation signal" for the rest of Subframe 0. We have more discussions on the reservation signal in the following subsection. In another case as shown in Figure 5 (b) where two systems are present, Cell A and Cell B are given the 19th and 5th CCA slots, respectively. Cell B listens on the channel during the 5th CCA slot before Cell A does (19th). Since no one is transmitting within this slot, Cell B wins the channel, and secures it by immediately transmitting the reservation signal for the remaining time of the subframe.

Since the finish of LBT does not guarantee the use right of the upcoming channel (in fact, the channel is up for grabs until RF energy is continuously radiated into the channel to secure the channel) and since the ending of the reservation signal most likely does not align with the subframe boundary,

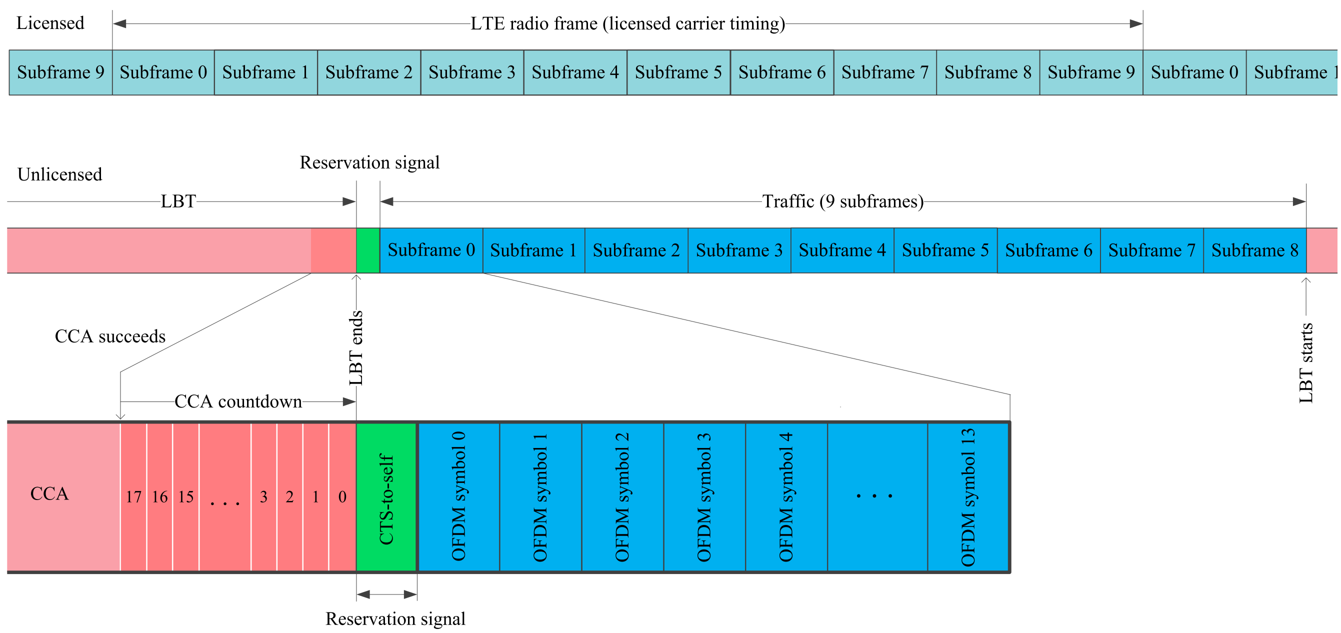


Fig. 6. Asynchronous load-based LBT, where frame timing is not synchronous to the system timing on licensed spectrum. CCA countdown starts immediately after the CCA is successful. The initial value of the counter is randomly selected from 0 to 31, and decrements as CCA continues to succeed.

the reservation signal is further extended to fill the gap till the traffic starts, e.g., the extended CRS shown in Figure 5.

On the other hand, Cell A does not start CCA until the beginning of the 19th CCA slot. Since Cell B is already transmitting (extended reservation signal), Cell A is thus presumably to fail the CCA. Cell A consequently yields to Cell B till the next LBT subframe.

This LBT structure may have solved the collision problem. However, another common issue associated with the frame-based structure is the potential blocking problem that may happen when cells are not synchronous. In that case, the CCA period from different systems may not be aligned, and one cell or cells may consistently block others. We revisit this issue in Section V-B *1) d)*.

2) Load-Based LBT: A realization of the LBT mechanism outlined in Figure 4 under the LTE frame structure is provided in Figure 6. During LBT, CCA is performed continuously without abiding any frame or OFDM symbol boundaries. Once CCA is successful, the CCA countdown starts with the counter initialized with a random number (17 within a contention window of 32 in this example). As soon as the counter counts down to zero, indicating the end of LBT, the reservation signal is transmitted to secure the right to use the channel for, e.g., 10 ms, subtracting the time consumed on the reservation signal.

A cell refrains itself from medium access until LBT ends. During LBT, CCA is continuously performed until it is successful. At this time, a random backoff timer is generated. The timer is decremented as long as the channel is idle (CCA is successful) but remains “frozen” when a transmission is detected (CCA fails), and reactivated as soon as the CCA succeeds. LBT ends when the timer expires. The random CCA countdown period provides additional random sensing time that helps avoid potential collisions which may happen when two or more transmitters are simultaneously waiting on the channel to be cleared.

Since the length of LBT is random, the ending of LBT is unpredictable and varies from time to time, which leads to a “floating” TTI that is asynchronous to the system operating on the licensed carrier. In other words, a floating subframe on the unlicensed carrier is likely misaligned with the boundary of the subframe on the licensed carrier, and the misalignment varies from one TTI to another. This timing uncertainty complicates the coordination between carriers, for instance, cross-carrier scheduling, HARQ timing determination, and CSI feedback timing. In addition, maintaining synchronism among carriers within the network to a common timing reference, i.e., the LTE carrier operating on the licensed band is desirable for coordination between these two types of carriers, which becomes clearer in Section V-B. In addition, synchronous LBT allows subframe boundary alignment across serving cells.

A means to synchronize the transmission timing to the system is illustrated in Figure 7. The synchronization is at the subframe level as well as the OFDM symbol level by design since resource scheduling in LTE is subframe-based. It is realized using the concept of dynamic LBT upon a fixed transmission framework to deal with the randomness in load-based LBT while still following a strict timeline synchronous to the system.

The radio frame is led by a special reservation signal, which is immediately transmitted to secure the channel as soon as the LBT finishes. For synchronization to the subframe, traffic does not start until the upcoming subframe starts. This may leave a gap between the reservation signal and the subframe [Subframe 3 in (a) and Subframe 4 in (b) of Figure 7 ranging from 0 (no gap) up to 13 OFDM symbols. Since the end of the reservation signal is likely

(a) Synchronous scenario 1

(b) Synchronous scenario 2

(c) Shortened backoff

Fig. 7. Load-based synchronous LBT, in which frame timing is synchronous to the anchor carrier on licensed spectrum. (a) Synchronous LBT scenario 1; (b) Synchronous scenario 2; (c) Shortened backoff mechanism is used to avoid consuming excessive time on reservation signal.

misaligned with the OFDM symbol timing, the reservation signal is *extended* to absorb the timing discrepancy just like the technique used in the frame-based LBT, allowing transition of the reservation signal timing (random) into the system OFDM symbol timing. The extension however is not fixed, but variable from zero OFDM symbol (inclusive) to one OFDM symbol (exclusive) depending on the ending point of the LBT.

C. Reservation Signal

As seen from the preceding two subsections, a reservation signal is transmitted before the start of traffic. It is used to secure the channel before the traffic can start in both frame-based and load-based LBT. The reservation signal essentially serves as "gap filler". For this purpose, any waveform will do the job as long as it satisfies the regulatory requirement, such as the 80% minimum bandwidth occupation (see Section II). However, the reservation signal may play another important role as a coexistence "coordinator" with the incumbent WiFi systems.

Each WiFi MAC frame contains a transmission duration field for informing the neighboring nodes of the medium occupancy time of the current burst. This is an amount of time that all nodes must wait if they receive it. A local timer, called network allocation vector (NAV), of a neighboring node is updated after the node reads the duration value from the ongoing transmission. This node defers from medium access until the NAV expires. Taking advantage of this "virtual medium sensing" mechanism, WiFi can simply use a special clear-to-send (CTS) message, referred to as "CTS-to-self" message for co-existence of new versions of WiFi frames with a legacy WiFi node. CTS-to-self is nothing but a standard WiFi CTS message in response to the ready-to-send (RTS) message that are originally created for dealing with the "hidden node" problem, except that here it is addressed to the transmitting node itself – as the name implies [36]. Nevertheless, it is meant for the neighboring WiFi nodes that are performing virtual medium sensing. A new generation WiFi node transmits a CTS-to-self frame right before transmitting, e.g., an 802.11n frame which is transparent to a legacy WiFi (e.g., 802.11b) node. The duration field of the CTS-to-self message contains the time of the following traffic frame (i.e., the 802.11n frame), thereby providing more effective protection of the subsequent frame(s) than that based on the physical medium sensing.

The same mechanism is adopted in this exemplary design, in which a reservation signal is led by the WiFi CTS-to-self signal [4]. Since the CTS-to-self is a WiFi message, it is understood and honored by all WiFi nodes that can receive it. As such, in the eyes of a WiFi system, a cellular system is no different than a regular WiFi system, and hence the medium access time as indicated in the CTS-to-self signal will be honored by a WiFi system, thereby providing better protection against WiFi transmissions. Since a cellular system in this design also honors the WiFi NAV mechanism during LBT, the protection naturally works both ways, ensuring smoother coexistence between cellular and WiFi systems.

As a design detail, since the CTS-to-self is a WiFi control signal, whose length (less than one OFDM symbol) is not flexible and not an integer multiples of an LTE OFDM symbol as shown in Figure 7 (a), the subsequent OFDM symbol is elongated by extending its cyclic prefix (CP) to align with the OFDM symbol boundary. Another issue associated with the inflexibility is that when the space between the end of LBT and the start of the upcoming subframe (Subframe 3) is less than the length of the CTS-to-self signal [Scenario 2 of Figure 7 (b)], the signal straddles between Subframes 2 and 3, and extends all the way to the start of Subframe 4. In this scenario, only 8 subframes are available for traffic, thereby consuming up to 13 OFDM symbols just for absorbing the timing difference between the LBT and the network timing. This overhead can be significant for the case when the allowed maximum channel occupation time is short. A way to avoid the excessive extension of the reservation signal in this scenario is to abort the backoff countdown prematurely to leave an extra space just enough to fit in the CTS signal before the next subframe starts, as depicted in Figure 7 (c). This early ending strategy may seem to unfairly disadvantage other competing systems. However, this shortened backoff can be compensated in the subsequent LBTs by voluntarily extending the CCA countdown time such that the overall average LBT time remains the same, so the fairness between different systems is strictly maintained.

D. Coexistence Analysis of LBT Structures

From the previous discussion, we see that LBT can be either load-based or frame-based. Frame-based LBT follows a fixed transmit/receive timing whereas load-based does not. In this sense, it seems that frame-based LBT would be a natural choice for LTE as LTE has a fixed transmission frame structure as well.

The problem is that it does not coexist well with systems that employ load-based LBT like WiFi since the fixed CCA window is disadvantageous in competing with the more aggressive load-based LBT that looks for opportunities basically all the time. We therefore devote this section to examining the performance of frame-based LBT in terms of channel occupancy time when coexisting with the load-based LBT.

From Figure 4, the behavior of a load-based system can be modelled as Figure 8 from the view of a frame-based system, where the load-based system is either in the "idle state", or "busy state", denoted as $\mathbf{I}$, and $\mathbf{I}_0$, respectively. The duration of an idle state is a variable multiples of an idle slot, τl, where τ is the length of the idle slot, and l a random variable that takes on a value from $\{1, 2, \cdots, L\}$. Here L corresponds to the backoff window of length τL of the load-based LBT.

The idle state is thus a collection of the idle periods with various lengths,

$$\mathbf{I} = \bigcup_{l=1}^{L} \mathbf{I}_l, \tag{1}$$

where

$$\mathbf{I}_l \triangleq \left\{ t_{\text{end}}^{(\mathbf{I})} - t_{\text{start}}^{(\mathbf{I})} = \tau l \right\}, \quad l = 1, 2, \cdots, L. \tag{2}$$

The duration of a busy state, $\mathbf{I}_0$, is equal to the length of a traffic frame, T of the load-based system, assumed to be fixed, e.g., $T = 10$ ms. Using the convention in [37], an idle slot and a traffic frame are collectively referred to as a virtual slot. Clearly,

$$\mathbf{I}_i \cap \mathbf{I}_j = \mathbf{\Phi}, i \neq j, \forall i, j \in \{0, 1, 2, \cdots, L\}, \tag{3}$$

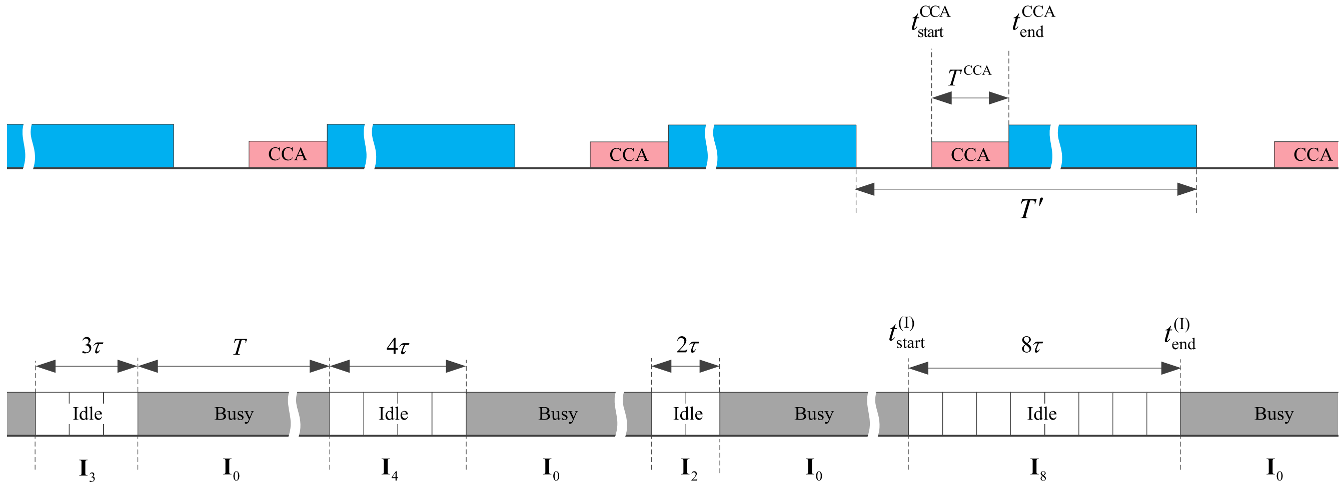


Fig. 8. Illustration of the channel state of a load-based system (bottom) from the view of a frame-based node (top).

and

$$P\{\mathbf{I}_0 \cup \mathbf{I}\} = P\left\{\mathbf{I}_0 \cup \bigcup_{l=1}^{L} \mathbf{I}_l\right\} = P\left\{\bigcup_{l=0}^{L} \mathbf{I}_l\right\} = 1, \tag{4}$$

where $\mathbf{\Phi}$ denotes the empty set.

From the perspective of a frame-based node, the channel status of a load-based system is thus a sequence of a mixture of $\mathbf{I}_l$, $l = 0, 1, 2, \cdots, L$. It is clear that a successful transmission from a frame-based node satisfies:

1) The start of CCA of the frame-based node, i.e., $t_{\text{start}}^{\text{CCA}}$, occurs in one of the idle periods, $\mathbf{I}$, of the load-based system,

$$\mathbf{\Gamma}_1 \triangleq \left\{t_{\text{start}}^{(\mathbf{I})} \leq t_{\text{start}}^{\text{CCA}} \leq t_{\text{end}}^{(\mathbf{I})}\right\}, \tag{5}$$

and

2) the end of the CCA period, $T_{\text{end}}^{\text{CCA}} = t_{\text{start}}^{\text{CCA}} + T^{\text{CCA}}$, is contained within the same idle period, $\mathbf{I}$,

$$\mathbf{\Gamma}_2 \triangleq \left\{t_{\text{start}}^{(\mathbf{I})} < t_{\text{start}}^{\text{CCA}} + T^{\text{CCA}} \leq t_{\text{end}}^{(\mathbf{I})}\right\}. \tag{6}$$

Both (5) and (6) must be satisfied in order for a frame-based node to succeed the CCA, meaning $\mathbf{\Gamma}_1 \cap \mathbf{\Gamma}_2 \neq \mathbf{\Phi}$. Using (3) and (4), the probability of $\mathbf{\Gamma}_1 \cap \mathbf{\Gamma}_2$ can be represented as

$$P\{\mathbf{\Gamma}_1 \cap \mathbf{\Gamma}_2\} = P\left\{\bigcup_{l=0}^{L} \mathbf{\Gamma}_1 \cap \mathbf{\Gamma}_2 \cap \mathbf{I}_l\right\} = \sum_{l=0}^{L} P\{\mathbf{\Gamma}_1 \cap \mathbf{\Gamma}_2 \cap \mathbf{I}_l\}. \tag{7}$$

Since $\mathbf{\Gamma}_1 \cap \mathbf{\Gamma}_2$ does not happen during the busy state,

$$P\{\mathbf{\Gamma}_1 \cap \mathbf{\Gamma}_2 \cap \mathbf{I}_0\} = 0, \tag{8}$$

Equation (7) becomes

$$P\{\mathbf{\Gamma}_1 \cap \mathbf{\Gamma}_2\} = \sum_{l=1}^{L} P\{\mathbf{\Gamma}_1 \cap \mathbf{\Gamma}_2 \cap \mathbf{I}_l\}. \tag{9}$$

And since the idle state must last longer than the CCA duration, $\tau l \geq T^{\text{CCA}}$ for CCA to succeed – a necessary condition for $\mathbf{\Gamma}_1 \cap \mathbf{\Gamma}_2$ to be true, $\mathbf{\Gamma}_1 \cap \mathbf{\Gamma}_2$ does not occur within $\mathbf{I}_l$ for $0 < l < L^{\text{CCA}} \triangleq \lceil \frac{T^{\text{CCA}}}{\tau} \rceil$ either. Equation (9) further reduces to

$$P\{\mathbf{\Gamma}_1 \cap \mathbf{\Gamma}_2\} = \sum_{l=L^{\text{CCA}}}^{L} P\{\mathbf{\Gamma}_1 \cap \mathbf{\Gamma}_2 \cap \mathbf{I}_l\}. \tag{10}$$

Notice that

$$P\{\mathbf{\Gamma}_1 \cap \mathbf{\Gamma}_2 \cap \mathbf{I}_l\} = P\{\mathbf{\Gamma}_2 | \mathbf{\Gamma}_1 \cap \mathbf{I}_l\} \cdot P\{\mathbf{\Gamma}_1 \cap \mathbf{I}_l\}, \tag{11}$$

where $P\{\mathbf{\Gamma}_2 | \mathbf{\Gamma}_1 \cap \mathbf{I}_l\}$ is the probability that the end of CCA falls within the idle period when so is the start of CCA. It is easy to verify that

$$P\{\mathbf{\Gamma}_2 | \mathbf{\Gamma}_1 \cap \mathbf{I}_l\} = \frac{1}{l}\left(l - T^{\text{CCA}} \big/ \tau\right), l \geq L^{\text{CCA}}, \tag{12}$$

and

$$P\{\mathbf{\Gamma}_1 \cap \mathbf{I}_l\} = P\{\mathbf{I}_l\} = \frac{(1-p) \cdot p^l \cdot l}{T \big/ \tau + (1-p) \cdot \sum_{i=1}^{L} p^i \cdot i}, \tag{13}$$

where p is the probability that the a virtual slot is idling in a load-based system. Substituting (11) – (13) into (10), we finally arrive at

$$\begin{aligned} &P\{\mathbf{\Gamma}_1 \cap \mathbf{\Gamma}_2\} \\ &= \sum_{l=L^{\text{CCA}}}^{L} \frac{(1-p) \cdot p^l \cdot l}{T \big/ \tau + (1-p) \cdot \sum_{i=1}^{L} p^i \cdot i} \cdot \frac{l - T^{\text{CCA}} \big/ \tau}{l} \\ &= \frac{1}{\frac{T}{\tau} + \left[\frac{p(1-p^L)}{1-p} - Lp^{L+1}\right]} \cdot \\ &\quad \left[\left(L^{\text{CCA}} + \frac{p}{1-p} - \frac{T^{\text{CCA}}}{\tau}\right)p^{L^{\text{CCA}}} \right. \\ &\quad \left. - \left(L + \frac{1}{1-p} - \frac{T^{\text{CCA}}}{\tau}\right)p^{L+1}\right]. \end{aligned} \tag{14}$$

Assuming that the frame period (i.e., idle period plus traffic period) of the frame-based node is T', the channel occupancy

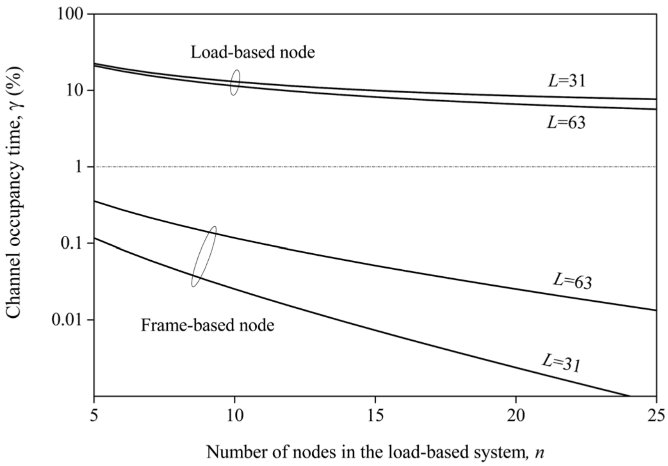


Fig. 9. Comparison of channel occupancy time between a frame-based node and a load-based node, where $T = 10$ms, $\tau = 9\mu$s, $L = 31$ and 63, $n = 10$ for the load-based system, and $T' = 0.5$ (idle) $+10$ (traffic) ms, $T^{\text{CCA}} = 25\mu$s for the frame-based system.

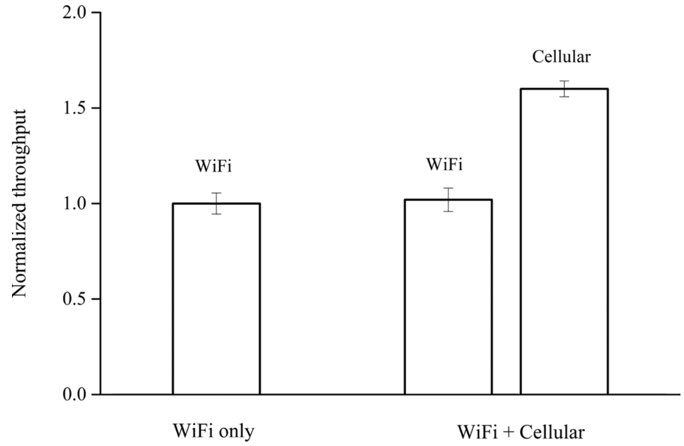


Fig. 10. WiFi and cellular on unlicensed spectrum system throughput (normalized to WiFi only scenario) to show the coexistence performance of the load-based LBT cellular system adopting the transmission mechanism described in Figure 7. The simulation scenario is based on Scenario 3 in [2] with the addition of unlicensed bands. 20 users per unlicensed carrier.

time of the frame-based node is

$$\begin{aligned}\gamma_{\text{frame}} &= \frac{T}{T'/P\{\boldsymbol{\Gamma}_1 \cap \boldsymbol{\Gamma}_2\}} \\ &= \frac{T/T'}{\frac{T}{\tau} + \frac{p}{1-p}\left(1-p^L\right) - Lp^{L+1}} \cdot \\ &\quad \left[\left(L^{\text{CCA}} + \frac{p}{1-p} - \frac{T^{\text{CCA}}}{\tau}\right)p^{L^{\text{CCA}}} \right. \\ &\quad \left. - \left(L + \frac{1}{1-p} - \frac{T^{\text{CCA}}}{\tau}\right)p^{L+1}\right]. \end{aligned} \tag{15}$$

And the channel occupancy time of a load-based node is

$$\begin{aligned}\gamma_{\text{load}} &= (1-\gamma_{\text{frame}})\frac{q \cdot T}{P\{\mathbf{I}_0\} \cdot T + \sum_{l=1}^{L} P\{\mathbf{I}_l\} \cdot l\tau} \\ &= (1-\gamma_{\text{frame}})\frac{qT/\tau}{(1-p)T/\tau + p - (1+L-pL)p^{L+1}}. \end{aligned} \tag{16}$$

Adopting the analytical results from previous studies on load-based systems [37], the probability that the system is idling during a virtual slot in a load-based system is

$$p = (1-q)^n, \tag{17}$$

where $q = \frac{2}{L+2}$ is the probability that a load-based node is transmitting in a virtual slot, and n is the number of nodes in the load-based system.

Equations (15) and (16) are plotted in Figure 9, where it is clearly seen that the frame-based LBT is disadvantageous in competing with load-based LBT, which is not surprising since load-based mechanism is more consistent and aggressive in terms of competing for resources. A load-based LBT mechanism is thus desirable in order to fairly coexist with load-based systems like WiFi.

E. Coexistence Simulations

A hybrid simulation is carried out to evaluate the coexistence performance between WiFi and the load-based LBT cellular system. A 2×20MHz unlicensed spectrum at 5 GHz is assumed in the simulation. Two scenarios are simulated. In the first scenario, 10 WiFi access points share the 40 MHz spectrum within a radius of 75 m. In the second scenario, 5 WiFi access points and 5 cellular small cells share the 40 MHz spectrum. Each cell employs the system described in Figure 7, operating on the two 20-MHz unlicensed bands shared with WiFi. The transmit power for both systems is 23 dBm. As shown in Figure 10, the WiFi system performance remains largely unaffected by sharing the spectrum with such a cellular system, which is expected due to (1) the use of a load-based LBT mechanism (which is essentially the WiFi DCF); and (2) the inclusion of the WiFi CTS-to-self signal as the reservation signal. It is not surprising to observe that the cellular system performs better than WiFi owing to the more efficient transmission structure and reliable control signaling provided by a licensed anchor carrier.

V. Practical Systems

A. LTE-U

The LTE-U Forum was formed by wireless industry leaders (Alcatel, Ericsson, Qualcomm, and Samsung) led by Verizon. The motivation was to facilitate carrier aggregation by way of unlicensed spectrum, exploiting the principles of cognitive radio to enable the LTE technology to utilize the vast amount of available spectrum in the less crowded 5 GHz band [38]. As depicted in Figure 1, there is up to 500 MHz of spectrum at 5 GHz.

In addition to advantages in channel coding, reference signal, and control channel designs, LTE possesses certain unique features that WiFi does not have. Indeed, LTE systems have a centralized scheduler that efficiently manages the use of spectrum to address both the overall network performance requirements and the specific needs of individual users. Particularly, it offers coordinated and synchronized scheduling of resources that can be distributed simultaneously among multiple users in both FDM and TDM fashions, whereas

WiFi tries to ensure that only one user is present in a channel at a time to avoid interference between users. Inter-cell interference in LTE is coordinated over the X2 interface, providing better cell-edge user experience and overall system capacity. In addition, LTE has more efficient link adaptation features such as HARQ for adapting to and recovering from channel errors.

However, direct deployment of LTE in the regions with the LBT requirements is prohibitive since a cellular system like LTE does not have the cognitive sensing capability. As a first step and for quick deployment, LTE-U Forum initially focuses on non-LBT markets without the need for modifying LTE PHY/MAC standards [38]. The deployment scenarios include LTE downlinks operating on unlicensed spectrum that contains multiples of contiguous 20 MHz bands. The LTE network with an uplink and downlink anchored on the licensed spectrum (less or equal to 20 MHz) for delivering critical control signaling [e.g., radio resource control (RRC) signaling] related to radio resource management (RRM) and connection/mobility provides reliability, mobility, and coverage, whereas the unlicensed spectrum is solely for the purpose of boosting the downlink data service. The "licensed-spectrum grade" control signaling provides additional significant performance gain of LTE-U over WiFi.

Referring back to Figure 1, an LTE-U deployment model is illustrated. The primary component carrier (anchor carrier) of the LTE-U is on the licensed spectrum employing the traditional LTE-Advanced or LTE-A whereas the unlicensed bands using LTE-U serve as secondary component carriers mainly for traffic transportation leveraging unlicensed spectrum to opportunistically offload the "best-effort" class of data traffic from the network. This configuration allows for exploitation of the ultra-wideband unlicensed spectrum for opportunistic and aggressive high rate data services while relying on the traditional cellular infrastructure on licensed spectrum for reliable control and high-QoS data services, as well as for coverage and mobility.

Without changes to PHY/MAC, there is still a need for a coexistence mechanism with other systems, particularly the incumbent WiFi systems, to maintain fairness (i.e., equal transmission opportunity) among different technologies. To this end, LTE-U employs cognitive radio principles to share the unlicensed spectrum with the incumbent WiFi systems. The spectrum is thus autonomously managed using cognitive sensing. Inspired by the cognitive radio concept, several medium access mechanisms have been proposed for fair coexistence with WiFi systems, such as virtual medium sensing, channel/carrier selection, adaptive muting, power control, and discontinuous transmission, etc., as briefly summarized below.

The dynamic channel selection mechanism supports cognition over the unlicensed bands, by dynamically selecting the channel with least interference to avoid using channels where nearby WiFi activity is detected. The 5 GHz band has more than 20 non-overlapping 20-MHz channels (see Fig. 1), which simplifies such channel selection. First, LTE-U looks for a cleanest channel where no evidence of WiFi activity is present by performing the WiFi specific measurement employing a WiFi network listening module that utilizes a WiFi specific detection scheme in addition to the universal energy detection CCA. To improve the sensing performance, it first detects the WiFi preamble. This waveform detection improves the detection sensitivity by $\sim$20 dB over the energy detection. If detected, the module tries to decode the following MAC frame that carries the payload. The duration field of the MAC payload indicates the medium occupancy time of the current transmission, based on which a local timer is updated if decoding is successful. Medium access is deferred from until the timer expires. This virtual medium sensing scheme clearly provides better protection against unwanted disruptions to the ongoing WiFi transmission than the *physical* medium sensing (i.e., the energy or waveform-based sensing). In the case of the radar channels, DFS must be applied to avoid channels with detected radar activities.

LTE-U monitors the status of the channel on an on-going basis, and selects and switches to a more suitable channel if needed. This carrier selection scheme is used to avoid co-channel operation with WiFi systems on a relatively slow time scale.

In the case when no clean channel is available, this algorithm shares the channel with WiFi systems following a TDM transmission pattern, comprised of ON-state and OFF-state. In the ON-state, LTE-U transmits according to LTE Release10 or later releases. In the OFF-state, LTE-U does nothing but sniffs the medium for co-channel WiFi activities, and adjusts the LTE-U duty cycle accordingly. This *adaptive muting* scheme is graphically illustrated in Figure 11. LTE-U transmits in ON-state and monitors the channel activities, particularly the WiFi activities, in OFF-state. The ON/OFF pattern or duty cycle changes over time based on the detected channel usage. In this example, due to the low WiFi activities detected during OFF-state, the following LTE-U duty cycle is increased accordingly. It is evident that an LTE-U system does not yield to the on-going WiFi traffic (if any). Instead, it follows its own transmission timing pattern that adapts to the WiFi traffic model derived from the sensing results on a much longer time scale than LBT. The coexistence performance thus highly depends on the accuracy of the model.

Adjustment of the transmission power might be used to assist coexistence by reducing interference, at the expense of reduced data rate [7]. It may even turn off the transmission in the unlicensed band, and use the licensed band if traffic load in the current cell is low.

In the event that channel selection and power control are not sufficient to avoid interference, LTE-U uses *discontinuous transmission*. For instance, during the ON-state, the LTE "*almost-blank-subframe*" (ABS) may be inserted (e.g., two blank subframes every 20 subframes) to provide opportunities for WiFi latency-sensitive applications to go through as shown in Figure 11.

However, for regions with the LBT regulatory requirement, the above cognitive radio based spectral management schemes are no longer sufficient. Transmissions must be controlled in a finer time scale. The LBT functionality operating on a much shorter time scale must be incorporated into the existing LTE transmission structure, which requires significant MAC and physical layer changes. LAA provides such a solution, which

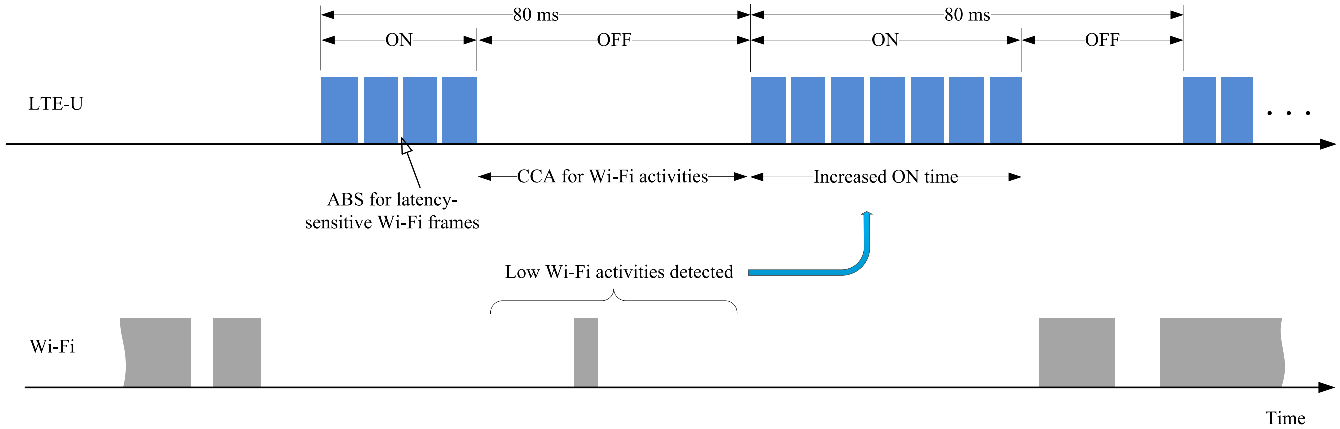


Fig. 11. Illustration of the LTE-U medium access mechanism to coexist with WiFi. In this example, due to the low WiFi activities detected during OFF-state, the following LTE-U duty cycle is increased accordingly. The gaps left in the ON-state are intended for flushing the WiFi time-critical frames.

is standardized in Releases 13 and 14 LTE, which is the topic of the next subsection.

B. LTE-LAA

In June 2015, 3GPP has approved a work item to standardize a technology that extends the LTE technology to the license-exempt spectrum to provide a universal framework with functionalities to meet the regulatory requirements in unlicensed spectrum, targeting a unified framework that complies with different regulations in different unlicensed bands and geographical regions [3], [38]. Another important goal is the fair coexistence with incumbent WiFi systems as well as among the new LTE systems deployed by different operators. Extension of the LTE downlink to the unlicensed band (e.g., the 5 GHz band) has recently finished as part of Release 13 LTE. The work is focused on "license-assisted access" or LAA on downlink, in which the access to unlicensed spectrum via a secondary component carrier is assisted by a primary component carrier on licensed spectrum (i.e., the anchor carrier) [39]–[42]. The corresponding uplink is to be included in Release 14.

In order to extend LTE to both non-LBT and LBT markets, changes must be introduced to LTE in order to meet the regulatory requirements on unlicensed spectrum in various regions, which are primarily on PHY/MAC. The upper layer (e.g., RRC) procedures may employ the existing LTE carrier aggregation framework, and control information is tunneled through the primary component carrier, i.e., the licensed LTE-A carrier. The new PHY/MAC features must also address the coexistence with WiFi systems as well as among LAA networks of different operators.

The choices of the three key feature schemes essentially determine the architecture of LTE-LAA: (1) The selection between frame-based LBT and load-based LBT structure. The choice of load-based LBT by LAA is justified by the analysis in Section IV. (2) The selection between synchronous and asynchronous LBT. The choice of synchronous LBT is also explained in detail in Section IV. (3) Should the reservation signal waveform be specified? In Section IV we briefly explained the functions of a reservation signal and more details on its merits and demerits are provided later in this subsection.

This section provides a detailed description on physical layer design issues based on the baseline framework established in Section IV.

1) Downlink: Two primary LBT schemes have been considered in LAA data transmission. They are: 1) LBT without random back-off which is basically the frame-based LBT (see Section IV-B); and 2) LBT with random back-off, at the same time, retaining the same transmission frame structure as the LTE licensed carrier, which is essentially the synchronous load-based LBT (see Section IV-B). LAA thus operates as a second carrier assisted by the primary carrier – the LTE licensed carrier.

From the analysis of the previous Section, it is not surprising that the load-based LBT is selected as the baseline LBT at least for LAA downlink transmission bursts containing PDSCH scheme for better compatibility with WiFi systems.

a) Reservation signal: As seen in Section IV-C, the reservation signal has the functions of preserving the channel to maintain synchronous timeline with LTE licensed carrier, and seamlessly coexisting with WiFi if the CTS-to-Self is used as part of the signal. The disadvantage is that the reservation signal consumes a part of the maximum channel occupancy time (MCOT), thereby leaving less for data transmission, which can be a significant overhead when the maximum occupancy time is small (e.g., 4 ms in Japan).

LAA does *not* specify the reservation signal waveform, leaving it out of the LAA transmission structure. Therefore the transmission of a reservation signal is completely up to implementation, which leaves the door open for future optimization. One may thus use the reservation signal as seen in Section IV-C for best compatibility with WiFi, or choose a waveform optimized for specific applications [43], [44]. One may even avoid using a reservation signal by extending CCA to align the traffic starting point with the subframe boundary as shown in Figure 12, which we henceforth refer to it as "self-defer", and the overhead is thus reduced (but at the cost of disadvantages in contending with other systems for channel usage). In addition, the absence of a *common* preamble may affect the reception performance as we will see in the following.

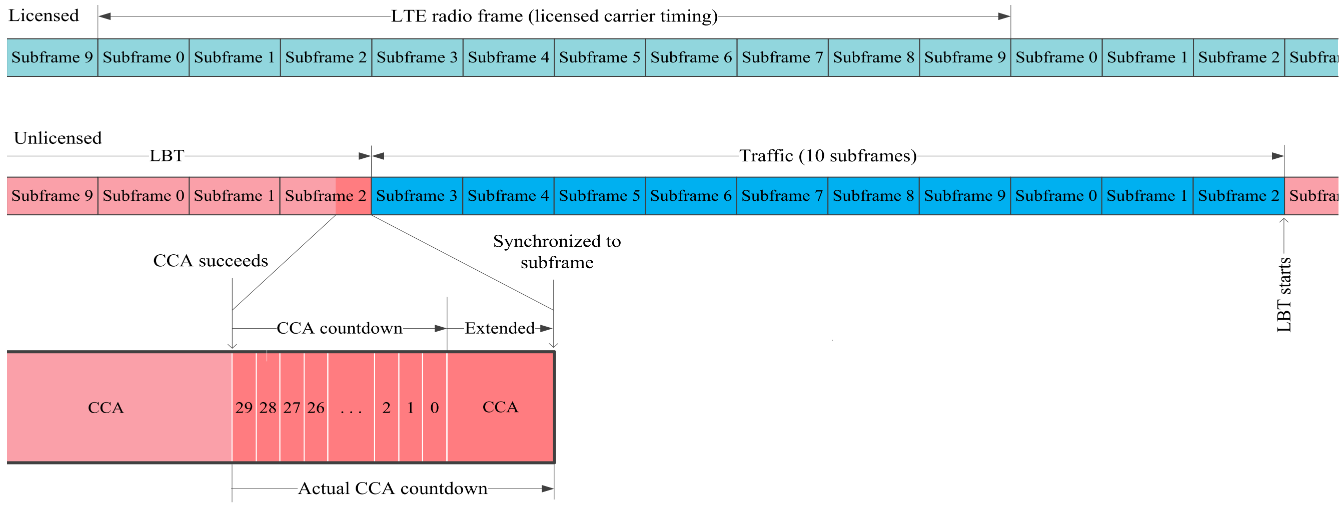


Fig. 12. A downlink transmission configuration without any reservation signal for maximum traffic time.

b) Partial subframe: From the previous examples, it is apparent that synchronous load-based LBT does not come without a cost. The overhead to maintain synchronousness, through either a reservation signal or self-defer, can be up to one subframe as shown in Figure 13 (a), which can be reduced if TTI can be a partial subframe. LAA thus allows for flexible data burst starting position in a subframe, leading to the introduction of “partial subframe” [45]–[56].

To reduce the medium time spent on synchronization to the subframe boundary, LAA allows the start of a traffic burst without abiding the subframe boundary. This results in the first data burst TTI being less than a subframe, while the subsequent TTIs get to be aligned with the subframe boundaries, keeping the synchronous nature of LAA TTI. Figure 13 (b) shows an example of the reservation signal occupying the channel until the next valid traffic start position in a subframe (the start of the second slot of subframe 2 in this example).

Being advantageous for better resource usage efficiency, partial starting subframe at the same time creates some practical issues: 1) The UE needs to detect each candidate starting OFDM symbols in the initial partial subframe; and 2) The LTE base station (or eNodeB) scheduler needs to prepare data in advance for all possible starting positions of downlink transport block in the initial partial subframe [57].

Detection of initial transmission: Since the reservation signal is not specified in LAA (i.e., no common preamble at the beginning of a transmission), a UE cannot assume the presence of a reservation signal, let alone the waveform of the reservation signal. Therefore, a UE cannot rely on the reservation signal for identifying the starting position of an initial downlink transmission burst. There are three potential detection alternatives [58]–[60]: 1) inserting an initial signal [e.g., primary synchronization signal (PSS)/secondary synchronization signal (SSS) from the legacy LTE] at the beginning of a data burst; 2) detecting PDCCH/EPDCCH; and 3) detecting the CRS signal within the subframe.

The first alternative relies on a special initial signal to facilitate reliable detection of the starting position of a partial subframe. However, this addition of a special signal beats the purpose of not specifying the reservation signal in the first place. (Remember that one of the reasons of not specifying a reservation signal is to give the implementer the option of saving overhead specially in the case of short maximum channel occupation time.) The second alternative would incur additional complexity for UE due to the need for blind decoding of control channels (PDCCH or EPDCCH) at any potential starting positions, whereas in the third alternative, the detection of the presence of CRS at any potential OFDM symbol boundary requires less complexity, but requires the starting position to be limited to the OFDM symbols with CRS. Although there are more than two OFDM symbols that contain CRS, taking into consideration the potential blind-decoding complexity, only the first OFDM symbol (i.e., the start of the first slot of a subframe) and the eighth OFDM symbol (i.e., the start of the second slot in a subframe) are selected as the potential starting positions. That is, the potential starting time of downlink transport block in the initial partial subframe is restricted to the slot boundary for the balance between complexity and performance. Compared with the legacy LTE, the difference is that control/data is also allowed to start in the second slot of the initial subframe.

Scheduling of the initial partial subframe: We now address the second issue associated with the initial partial subframe. We start by a brief review of how a cellular system like LTE manages the medium access.

Resources in a cellular system, i.e., the licensed spectrum, are strictly managed by the network that owns the spectrum. This centralized multiple access method known as centralized “scheduling” allows transmissions with collision protection and issues like fairness, QoS, channel condition to be taken into consideration when allocating resources among users. In LTE, eNodeB uses the downlink grant to inform each UE on which downlink resources (i.e., the physical downlink shared channel or PDSCH). As mentioned in Section IV-A, both the downlink grants and uplink assignments are transmitted through the downlink control channel (i.e., PDCCH or

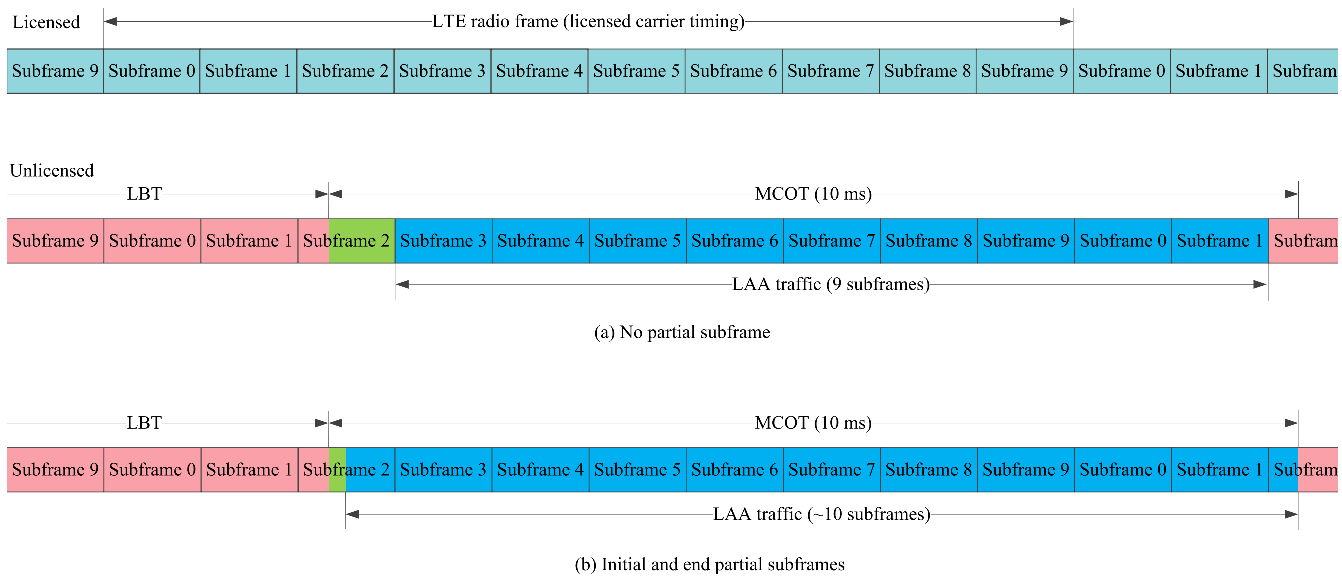


Fig. 13. a) LAA without partial subframe; b) LAA with partial subframe mechanism.

EPDCCH) on a subframe basis (i.e., the TTI is a subframe). Since scheduling is carried out per subframe, the UE only needs to buffer one subframe to obtain the potential scheduling information on PDCCH (EPDCCH) and use it to decode the associated PDSCH in the *current* subframe.

Like in LTE, LAA resources (the resources that are claimed by LAA after LBT) are controlled by the network. Clearly, if a control message regarding the data transmission (e.g., grant) is not received, the corresponding data transmission will also fail. Consequently, the control channels must be designed with sufficient reliability so that the control message is correctly received anywhere within a cell. Moreover, unlike the transmission of the user data, PDCCH does not have the luxury of HARQ since it itself requires the use of control channels, implying that a UE only has one chance to decode its PDCCH. Unfortunately, reliability is hard to guarantee on unlicensed spectrum as different systems compete for resources.

We recall that on licensed spectrum, only one system (the system that owns the spectrum) is allowed to use the spectrum. As such, a scheduler can continuously utilize the resources and efficiently manage them freely without the need for monitoring the channel activities or yielding to the traffic from other systems. The "assistance" from a licensed LTE primary carrier is thus extremely beneficial for conveying control messages between the network and mobile users via, e.g., the downlink control channel, e.g., PDCCH on the licensed carrier in a *timely* and *reliable* fashion. Therefore, all control signaling (from physical layer as well as higher layers) is preferably delivered via licensed carrier, just as the name LAA implies.

However, with the introduction of partial initial subframes, the traditional subframe-based scheduling scheme needs to be revised to take into account the new TTI boundary.

When the LBT finishes within the second slot of the subframe, the traffic starts in the following subframe. For this kind of regular full subframe TTI, on-carrier scheduling based on either local PDCCH or EPDCCH is supported in LAA, just like in the legacy LTE. Besides, cross-carrier scheduling from licensed carrier to unlicensed carrier using either PDCCH or EPDCCH is supported as well for better reliability. However, cross-carrier scheduling between unlicensed carriers is not supported (not to mention the scenario of using an unlicensed carrier to schedule a licensed carrier). A UE can expect the related LAA PDSCH to be present in case of receiving a valid respective grant from PDCCH/EPDCCH in the current subframe.

When the time that the channel becomes available (i.e., the end of LBT) passes the end of the control region but is within the first slot of the subframe (i.e., the second slot is available for traffic), it is not possible to transmit the cross-carrier scheduling information via either PDCCH or EPDCCH on the licensed carrier in this subframe as shown in Figure 14, as the eNodeB cannot predict when the LBT finishes. A straightforward solution is simply to defer the transmission of the scheduling information to the next subframe, i.e., the grant for the current (partial) subframe is sent on the next (full) subframe. This requires the UE to buffer all the data symbols before the grant is successfully decoded in the next subframe.

We may also use the PDCCH or EPDCCH on the current (unlicensed) carrier to schedule the current subframe on the current carrier. In this case, the design of the legacy LTE PDCCH for a full subframe can be reused for LAA partial subframes by shifting the PDCCH to the second slot of the initial partial subframe (as depicted in Figure 14). The number of PDCCH OFDM symbols is signaled using the PCFICH channel (on the current carrier) as in the legacy LTE. Similarly, if the EPDCCH is used for on-carrier scheduling of the partial subframe, it is offset by 7 OFDM symbols relative to the legacy EPDCCH [22]. Resource mapping follows the same rule as for the regular full subframe except that the number of available REs for the EPDCCH is halved, and higher aggregation level [61] may be needed to make up for the dimension loss. Nevertheless, the coverage of the EPDCCH is to some degree compromised.

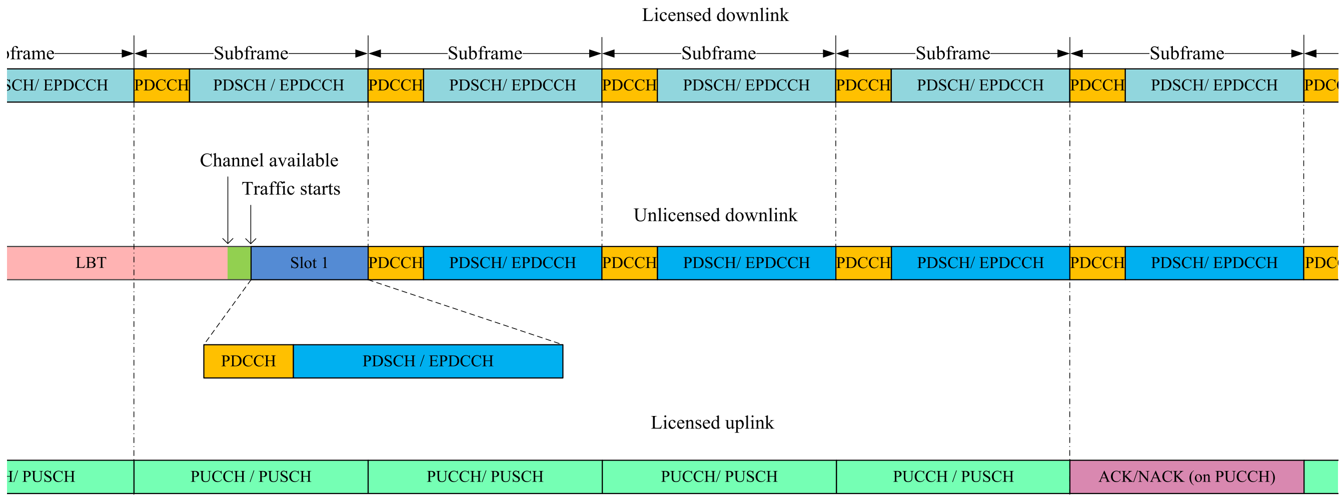


Fig. 14. Timing association between licensed control channel and unlicensed traffic. When the end of LBT passes the end of the control region but is within the first slot of the subframe (Slot 0), it is thus not possible to do cross-carrier scheduling via either PDCCH or EPDCCH on the licensed carrier in this subframe. The UE is expected to blindly decode PDCCH/EPDCCH on unlicensed carrier in the second slot. If scheduled, the corresponding ACK/NACK is fed back on the licensed uplink control channel (PUCCH). PUSCH denotes the physical uplink shared channel.

The deferred cross-carrier scheduling for partial subframes has the merit of requiring minimum change in the downlink control channel, as well as no need for blind control channel decoding in the second slot. Nevertheless, it does require up to one subframe additional buffering at the UE. On-carrier scheduling on the other hand does not require additional buffering, and hence is selected by LAA, although it involves more specification efforts.

The expected UE behavior is summarized in Figure 15. At the subframe boundary, a UE blind-decodes the PDCCH/EPDCCH on the licensed carrier for possible scheduled PDSCH (on the subframe basis) as for the legacy LTE UE. If a PDSCH on the unlicensed carrier is scheduled, the UE proceeds to decode the PDSCH. If no PDSCH is scheduled, the UE detects the CRS on the unlicensed carrier. If detected, the UE blind-decodes the PDCCH and EPDCCH for scheduled PDSCH on unlicensed carrier and proceeds accordingly. If the CRS is not detected, a UE capable of handling partial subframe will perform the same procedure in the second slot of the subframe, while a UE not capable of processing partial subframe will skip the procedure.

It is worth noting that, since the reservation signal may start at any time within a subframe, the last TTI within an MCOT is likely shorter than a subframe as illustrated in Figure 13 (b). This scenario is readily dealt with via what so called downlink pilot time slot (DwPTS) structure in the legacy LTE [62]. The applicable DwPTS configurations include 3, 6, 9, 10, 11, 12 OFDM symbols, of which a UE needs to be informed. Hence LAA adds an ending subframe configuration signaling in the common search space of control channel to indicate to the UE if the ending subframe is a full one or a DwPTS partial subframe configuration.

c) Hybrid ARQ: As the uplink control channel PUCCH is not supported in unlicensed uplink either in Release 13 (or in Release 14), the ACK/NACK feedback, in association with an unlicensed downlink traffic channel PDSCH, is conveyed through the more reliable PUCCH of the licensed carrier as depicted in Figure 14. The HARQ process for unlicensed downlink remains to be asynchronous as in the licensed downlink (see Section IV-A).

d) Discovery signal: Discovery reference signals (DRS) with a low duty cycle was originally introduced in small cell networks to provide the functionality of synchronization, cell identification, and RRM measurement for UEs when a small cell is in the dormant state [2]. The network makes a decision on whether to activate the small cell based on the RRM measurement report. The DRS waveform is primarily a combination of PSS, SSS, and CRS.

Functions of DRS: Generally, an LAA UE may use the licensed frame structure timeline for LAA in collocated scenario (i.e., LAA and licensed LTE cells are situated in the same cell cite). However, in non-collocated scenarios, due to the propagation delay difference, the timing discrepancy between the LTE licensed cell and LAA cell can be as large as tens of microseconds. In this respect, an LAA DRS provides a potentially better means for synchronizing LAA UEs.

In addition to the benefit of providing a means for synchronization, DRS can be used for signal quality measurements. When the measured values, e.g., reference signal received power (RSRP) or reference signal receiving quality (RSRQ) exceed a certain threshold, a measurement report from the UE will be sent to the eNodeB via licensed carrier to assist cell selection, carrier selection, and radio resource assignment on unlicensed carriers [21].

LBT for DRS: Unlike the traffic burst, DRS transmission in LAA adopts a transmission structure similar to the frame-based LBT (see Figure 3), resulting in a periodic DRS transmission to satisfy the more stringent timing requirement. As illustrated in Figure 16 (a) the DRS is set to be transmitted

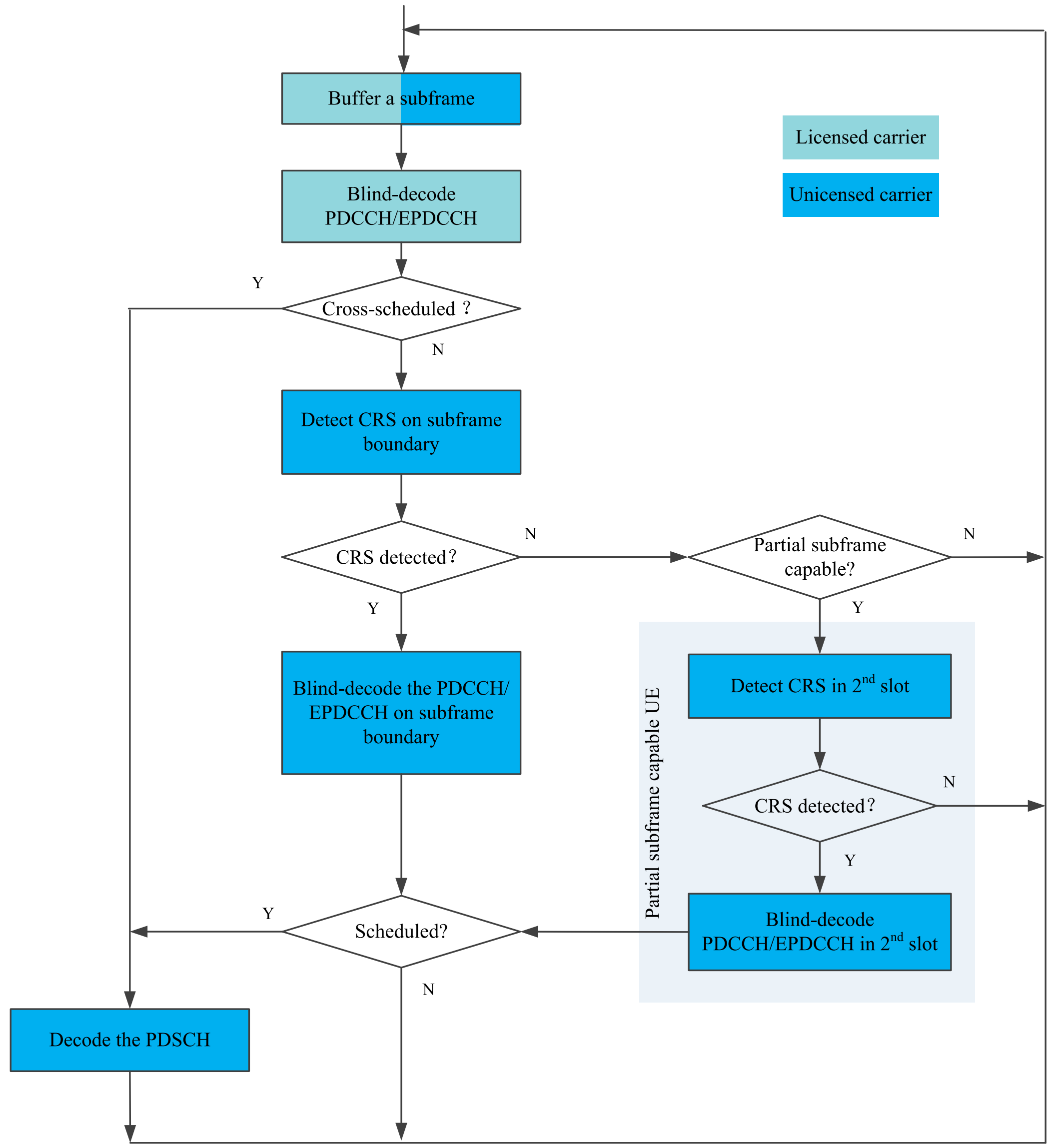
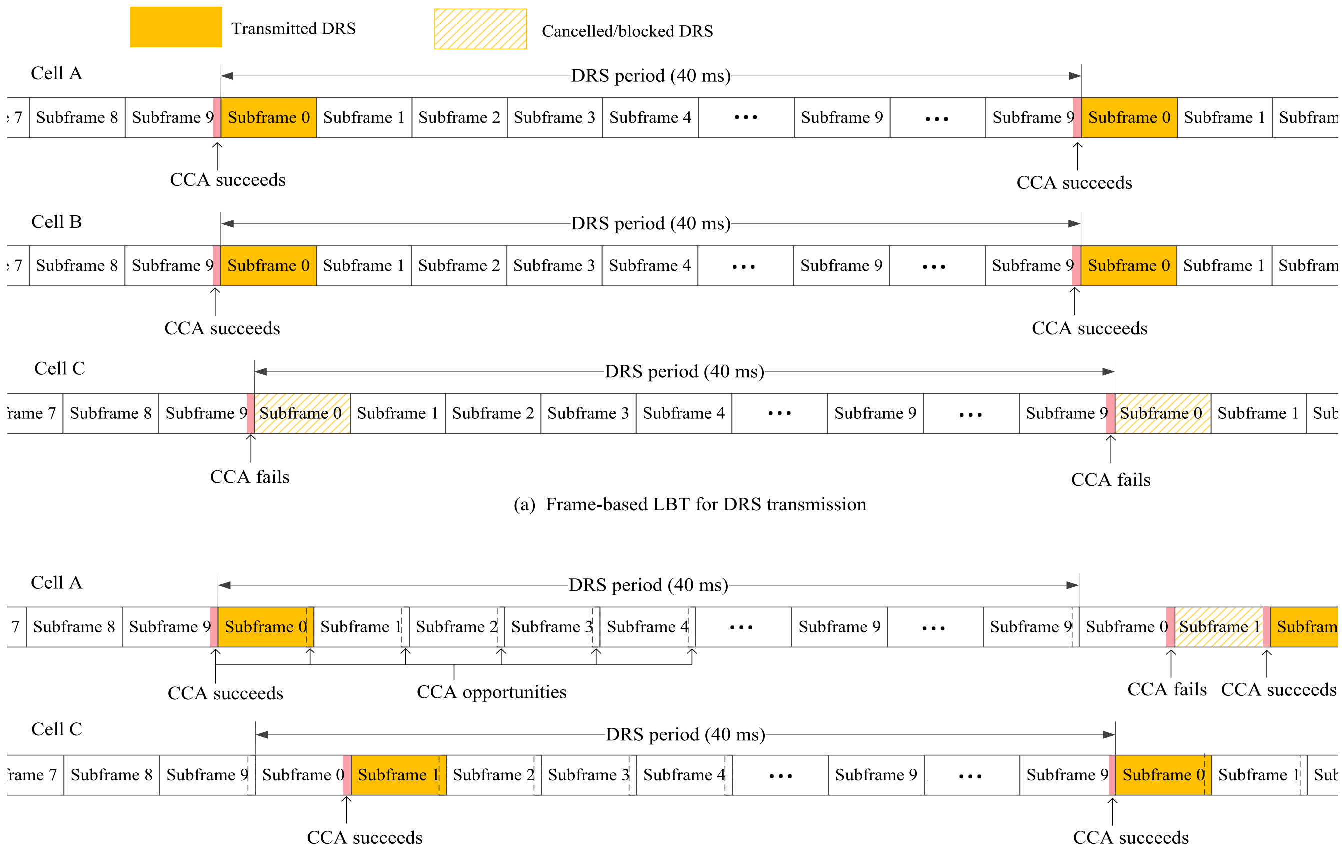


Fig. 15. Expected LAA UE behavior.

in subframe 0 every four radio frames (40 ms). A UE monitors DRS in, e.g., subframe 0 every 40ms based on the timing from the licensed carrier. Apparently, the transmission of DRS is opportunistic as a result of the CCA.

As pointed out earlier in Section IV-B1), a potential blocking problem associated with the frame-based LBT may arise when cells (e.g., from different networks) are not synchronized (which may very well be the case in practice). As shown in Figure 16 (a), the DRS CCA of cell C is within the DRS transmission duration of cell A, and hence cell C DRS is consistently blocked. A direct solution to preventing this monopoly is to provide a cell with more DRS transmission opportunities. As shown in Figure 16 (b), five consecutive subframes (subframes 0, 1, 2, 3, 4 for instance) rather than a single subframe, are used for potential DRS transmissions. An eNodeB may randomly select a subframe and perform CCA or simply go through each CCA opportunity until CCA succeeds or all CCA opportunities are exhausted. This DRS LBT mechanism increases the DRS transmission possibilities and alleviates the blocking issue at the cost of increased complexity of DRS detection. However, in either case, the transmission of DRS is not guaranteed.

2) Uplink: The main challenges of the unlicensed uplink are: (1) The regulation requires that the occupied transmission bandwidth for a transmitter be more than 80% of the total bandwidth (see Section II). This requires the change of the existing LTE uplink single-carrier FDM (SC-FDM) waveform, whereas for the downlink OFDM waveform this problem can be easily avoided. (2) The coordination between scheduling and LBT. The coordination is difficult for uplink since scheduling of uplink transmissions (by eNodeB) and LBT (by individual UE) are non-collocated, whereas for downlink they are collocated (at the eNodeB). R13 focused on the downlink first since the implementation of downlink is relatively simpler than uplink and boosting of downlink capacity is more beneficial as traffic is typically downlink dominated.



(a) Frame-based LBT for DRS transmission

(b) Modified frame-based LBT for DRS transmission

Fig. 16. Illustration of LAA discovery signal transmission. (a) The blocking issue happens among asynchronous cells (e.g., from different operators) when the frame-based LBT mechanism is employed for DRS transmission; and (b) the modified frame-based LBT mechanism that mitigates the blocking issue through the use of more DRS transmission opportunities within a DRS period.

In this subsection, we describe the uplink multiple access scheme in unlicensed band. We first look at the impact of the regulatory requirements of unlicensed spectrum on the uplink multiple access waveform.

a) Access waveform: In legacy LTE, uplink transmission utilizes SC-FDM to deal with the assignment of multiple UEs in an FDM manner on licensed spectrum. SC-FDM is an OFDM modulation scheme that emulates a single-carrier waveform with the benefit of reduced peak-to-average power ratio (PAPR). In OFDM, N modulation symbols are transmitted in parallel (not necessarily contiguous), one per subcarrier (15 kHz) with a duration of one OFDM symbol, whereas in SC-FDM N modulation symbols are transmitted sequentially, each occupying the bandwidth of N *contiguous* subcarriers. OFDM waveform is basically multi-carrier in nature, and SC-FDM more like single-carrier. In LTE, the granularity for uplink resource allocation is one RB (see Section IV-A), and a PUSCH contains a set of *contiguous* RBs, giving rise to a carrier spanning an integer multiple of 15 kHz continuous bandwidth as depicted in Figure 17 (a). For a single-carrier waveform, the PAPR is the PAPR of the modulation symbol, whereas for multi-carrier, the time-domain waveform is the sum of multiple modulation symbols on multiple subcarriers, creating higher PAPR.

Recalling that the LBT regulation requires the occupied transmission bandwidth to be at least 80% of the total bandwidth and with a maximum power spectral density (PSD), e.g., 10 dBm per MHz (see Section II), LAA thus introduces *cluster* as the base unit for unlicensed uplink resource allocation, which consists of 10 RBs evenly-distributed over 100 RBs in a 20 MHz bandwidth as shown in Figure 17 (b). A PUSCH may include multiple clusters with total transmission power up-bounded by the maximum PSD per MHz. Figure 17 (b) shows two examples of PUSCH resource allocation for four UEs in an FDM manner, where each UE is allotted one cluster in the first example, while UE B is assigned two in the second example. The drawback of the cluster-based waveform is the increased PAPR as compared to the traditional SC-FDMA waveform due to the loss of the single-carrier property.

b) Multiple access: As pointed out in Section III-B, in a centralized multiple access network, the utilization of the resource is fully controlled by the "scheduler" located in eNodeB, whereas, for operating on the unlicensed spectrum, scheduling must obey LBT. The coordination of scheduling and LBT is relatively easy to do for downlink since LBT is also performed at eNodeB. It is more difficult for uplink since the LBT is done at each individual UE and these LBT results

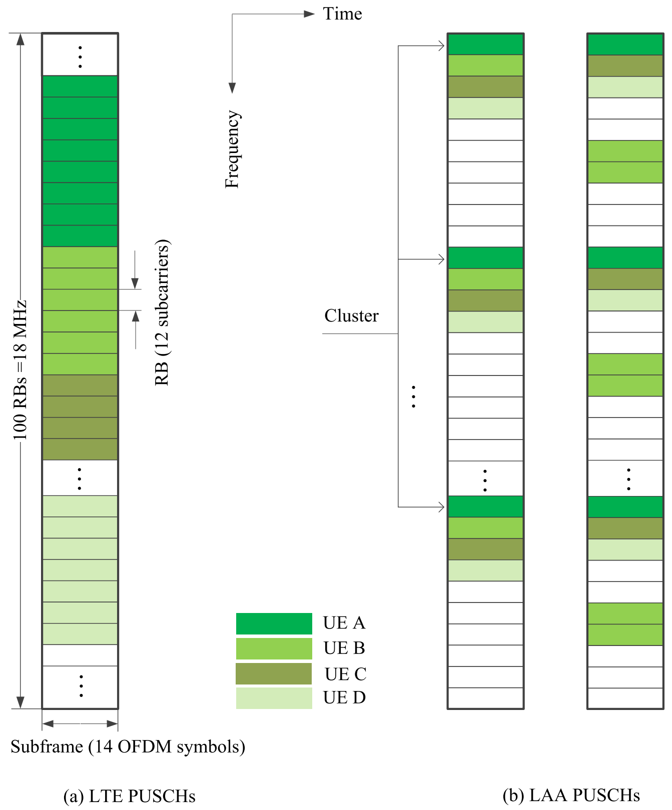


Fig. 17. Examples of PUSCH resource allocation for four UEs. (a) SC-FDM PUSCH resource allocation; and (b) LAA PUSCH resource unit, a cluster. A cluster consists of 10 RBs evenly distributed over 100 RBs in a 20MHz unlicensed band.

are not available at the eNodeB when scheduling uplink transmissions. Moreover, unlike in WiFi, where one transmission at a time is targeted by DCF, LAA supports simultaneous transmissions among multiple UEs in an FDM fashion to improve the resource usage efficiency just like in the legacy LTE, which further complicates the uplink transmissions.

Parallel UE LBT multiple access: In this access scheme, scheduling is independent on the LBT performed by the UE. An LBT step is thus imposed on each *scheduled* UE before transmission. Therefore, there is no guarantee for the scheduled UE to complete the LBT before the scheduled transmission time due to the uncertainty of LBT. As is representatively illustrated in Figure 18, UE (which are UE A, UE B, and UE C) starts LBT after an uplink assignment is received (from the licensed carrier control channel, either PDCCH or EPDCCH). Depending on the *local* interference, some of the UE (UE A and UE B) may be able to finish the LBT before the scheduled time (Subframe 4), and defer their transmissions accordingly till the indicated time, while those who are not able to finish LBT (UE C) must cancel the scheduled transmissions. During the deferring period, UE continues CCA and refrains itself from any kind of transmission (not even the reservation signal) which otherwise may block other scheduled UE who are performing CCA, henceforth referred to as "self-blocking".

Note that in the uplink, the timing advance (as depicted in Figure 18) is used by the eNodeB to advance each UE's uplink timing to account for its propagation delay such that signals from different UEs are aligned at the eNodeB receiver. The CP of the OFDM symbol thus only needs to accommodate the multipath delay spread of UE. Evidently, depending on the distance to the eNodeB, different UEs may have different timing advances. It may thus seem that the scheduled transmission of a UE (e.g., UE A) may interfere with another UE's CCA (e.g., UE B). However, time-advanced transmissions from different UE do not cause self-blocking issues among scheduled UE. Although UE A transmits ahead of UE B by $t_{\mathrm{Ae}} - t_{\mathrm{Be}}$, by the time UE B hears the transmission from UE A (which takes t_{BA} to reach UE B), it has already finished CCA and started transmission as well, noting that

$$t_{\mathrm{BA}} + t_{\mathrm{Be}} \geq t_{\mathrm{Ae}} \text{ or } t_{\mathrm{BA}} \geq t_{\mathrm{Ae}} - t_{\mathrm{Be}}, \tag{18}$$

per *the triangular inequality* (see Figure 19). This property allows FDM among UEs with timing advances.

The merit of this uplink transmission scheme is that each successful LBT by scheduled UE warrants the UE a *full* MCOT worth of transmission duration. Nonetheless, the demerit is that the difficulty of coordination between network scheduling and UE LBT results in cancelled transmissions (UE C in the example of Figure 18). Moreover, the full-fledged LBT may create a large vacuum between the end of LBT and the beginning of the subframe, leaving it vulnerable to potential disruptions from other systems since no reservation signal can be transmitted during that period due to self-blocking.

Shared MCOT multiple access: A solution to the problem in the parallel LBT access is simply to leave the LBT to the eNodeB, i.e., LBT and scheduling are performed jointly at the eNodeB. An alternative scheme for LAA uplink transmission is thus that after a successful LBT at the eNodeB, the MCOT is shared between downlink and uplink, i.e., TDD between downlink and uplink transmissions is carried out within the MCOT, thereby negating the need for LBT by UE. Consequently, uplink resource scheduling and transmissions following the downlink can be done in the same way as in the licensed LTE uplink, as depicted in Figure 20.

An immediate problem with this scheme is associated with the uplink scheduling delay. Thinking back to the necessity of scheduling delay (four subframes) mentioned in Section IV-A, an uplink transmission cannot start until the 5^{th} subframe of the MCOT. Referring to the example of Figure 20, the scheduling information carried by PDCCH is transmitted at least 4 subframes ahead of the transmission of the associated PUSCH, and the uplink assignments cannot be issued until the LBT is completed (remembering that the time that LBT ends is not predictable). This leaves the maximum amount of medium time available for PUSCH to four subframes in the case of 8-ms MCOT. This transmission scheme is thus not suitable for short MCOTs, and fails to work in regions with a regulated MCOT shorter than 5 ms (e.g., 4 ms in Japan).

This difficulty of enforcing centralized scheduling in unlicensed spectrum stems from the random nature of communications on the unlicensed spectrum that voids the predictability presumption by scheduling. Nevertheless, this problem can be ameliorated by means of a *two-stage assignment*. A two-stage assignment consists of an *initial uplink assignment* specific

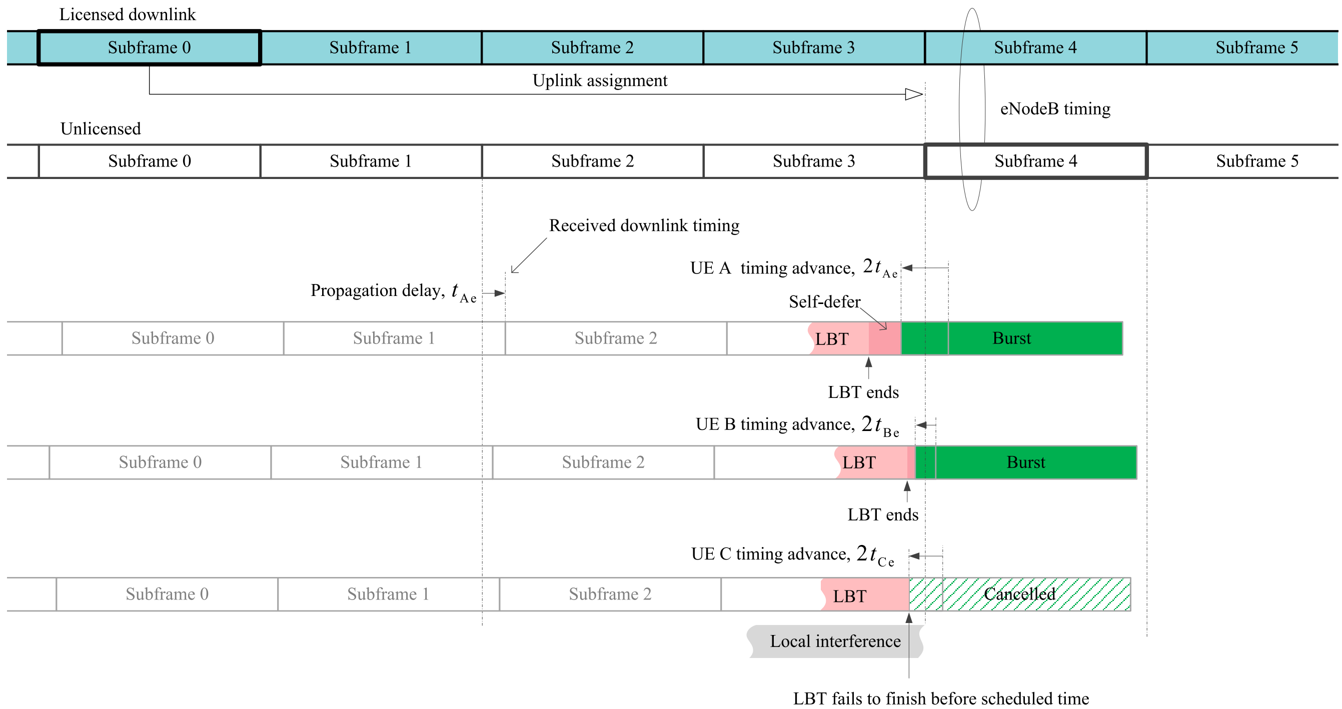


Fig. 18. Illustration of "parallel UE LBT multiple access". Unlicensed timeline is aligned with the licensed carrier timeline. Each UE performs timing advance to align with the timeline at the eNodeB, where t_{Ae}, t_{Be}, and t_{Ce} are the propagation delay from eNodeB to UE A, UE B, and UE C, respectively. The hollow arrow indicates the uplink assignments on the unlicensed carrier for UE A, UE B and UE C. The scheduling delay is 4 subframes.

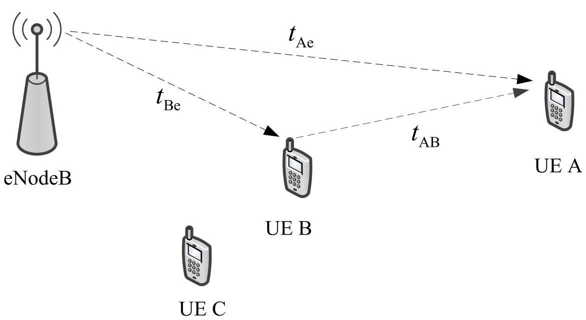


Fig. 19. Illustration of propagation delay in Figure 18, where t_{YX} denotes the propagation time from X to Y. For instance, t_{Ae} denotes propagation time from eNodeB to UE A, and t_{AB} from UE B to UE A.

to a UE, issued before the end of LBT, and a *trigger* signal common to all UEs (that are scheduled in the initial assignments), issued upon the finish of LBT, which indicates the start of the PUSCH region and the duration of this region (Figure 21). The inclusion of the starting point (or time offset) of the PUSCH region in the trigger signal gives the scheduler more flexibility to adjust the partition of downlink and uplink medium time to balance the downlink and uplink traffic dynamically. In addition to the scheduling information as included in the traditional uplink assignment, the initial assignment also includes the UE's PUSCH position in the PUSCH region, and the search window for the trigger signal.

With successful decoding of the initial assignment, the UE, if scheduled in the assignment, has sufficient information for preparation of the data for transmission, and starts transmitting as soon as the trigger signal is received. This two-stage assignment scheme is illustrated in Figure 21.

Although LBT performed by eNodeB negates the need for LBT at the UE, considering the different interference environment at the eNodeB and at the scheduled UE, CCA at UE is necessary before the scheduled transmission to avoid interfering with potential ongoing transmissions (from other systems) near the UE, which may not be detected by the eNodeB. A quick CCA check (e.g., 25 µs) is hence performed by UE before transmission. To this end, the first OFDM symbol ($\sim 70\mu$s) at the beginning of the PUSCH region is set aside as illustrated in Figure 22, meaning that eNodeB expects the uplink data to start from the second OFDM symbol. This creates one OFDM symbol worth of a gap between the downlink and uplink.

Noticing from Figure 22, timing advance is used by the eNodeB to advance each UE's uplink timing to account for its round-trip propagation delay to ensure that signals from different UEs are aligned together and synchronized to the system timing at the eNodeB receiver such that downlink and uplink have the same timing (subframe and OFDM symbol) at eNodeB [63]. Evidently, depending on the distance to the eNodeB, different UEs may experience different propagation delays, and hence have different timing advances.

Consequently, as seen from Figure 22, the actual length of the gap seen by a UE depends on the propagation delay and timing advance. It thus varies from UE to UE. Nonetheless, the minimum gap must be sufficiently large to accommodate a quick CCA. The maximum propagation delay can be supported by an OFDM symbol is therefore $(70\ \mu s - 25\ \mu s)/2 \approx$

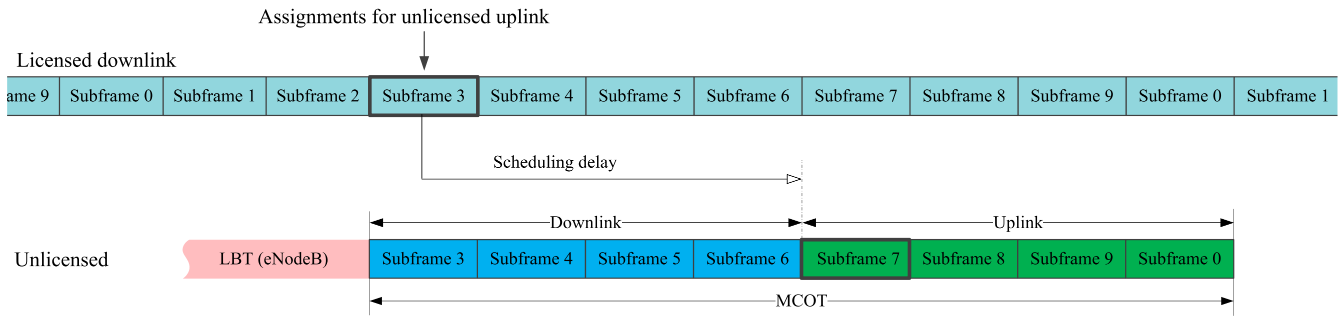


Fig. 20. Downlink and uplink-shared MCOT, in which no LBT is performed for unlicensed uplink transmission.

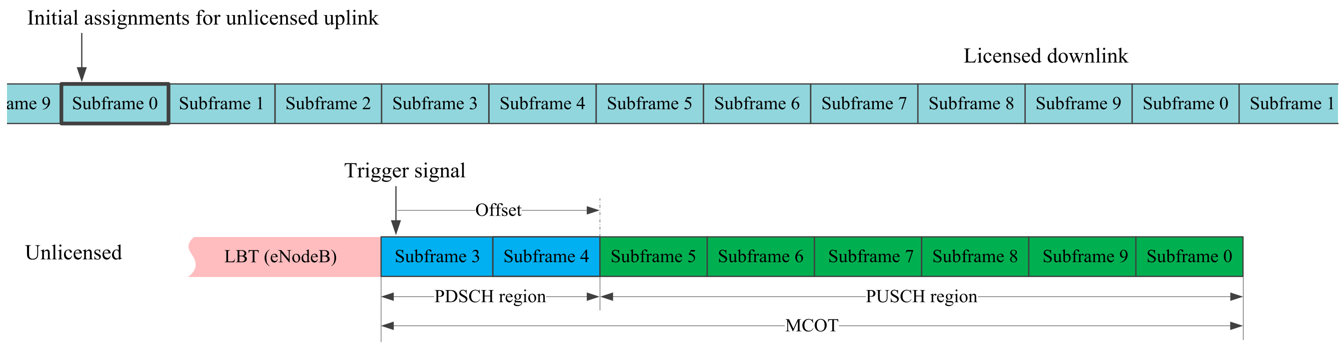


Fig. 21. A two-stage assignment for uplink scheduling, in which an initial assignment allows UE to prepare data for transmission, and a trigger signal to indicate the actual uplink starting time and duration. The value (ranging from one subframe to six subframes) of the offset included in the trigger signal can be controlled by the scheduler per the MCOT and downlink and uplink traffic loads. The duration can be up to six subframes. Here, only the timing at subframe level is shown, whereas the timing at a finer resolution is deferred to Figure 22.

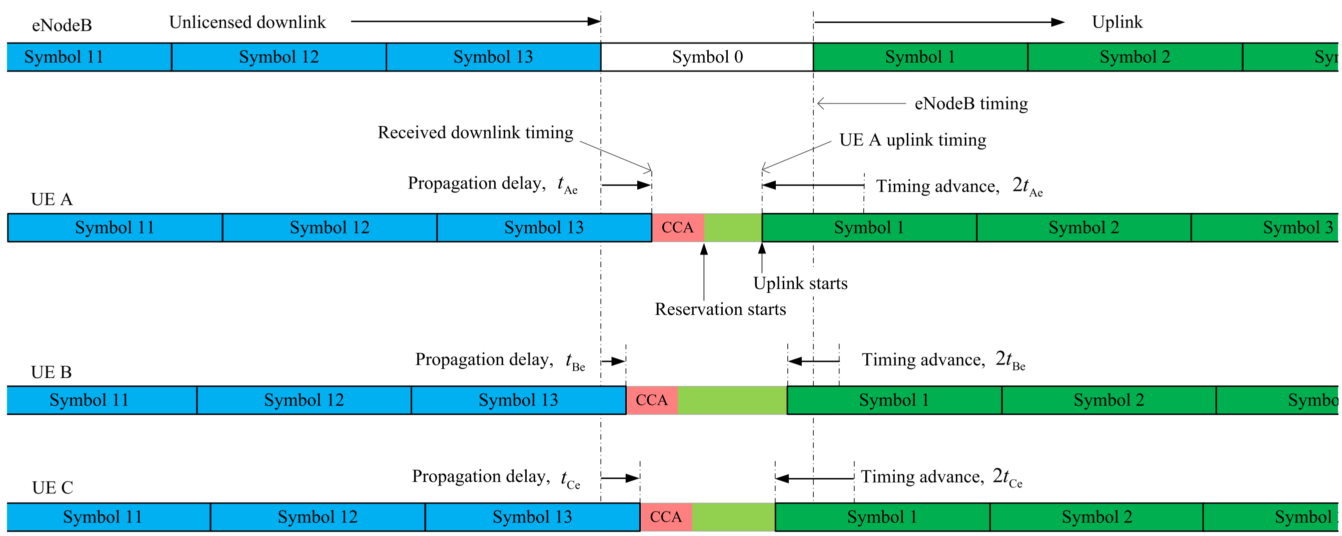


Fig. 22. The first OFDM symbol at the beginning of the uplink transmission is set aside as the "gap" at the transition from downlink to uplink. The gap duration varies per UE due to the propagation delay and timing advance. A reservation signal is inserted to secure the channel as soon as CCA succeeds.

23μs, corresponding to a cell radius of 7 km, which is sufficient for a small cell.

Utilizing this gap, a scheduled UE performs a CCA to check for any local ongoing transmission as soon as the downlink transmission ends. A reservation signal is then inserted all the way to the start of the second OFDM symbol when the uplink traffic is supposed to start. Here, this reservation signal is simply the extension of the CP of the second OFDM symbol.

However, it seems that, due to different propagation delays, the reservation signals from some scheduled UEs may unintentionally silence other scheduled UEs that are performing CCA. Taking Figure 22 as example, the reservation signal transmitted by UE B after CCA may seem to fail the CCA that UE A is performing. However, since

$$t_{\mathrm{Be}} + t_{\mathrm{AB}} \geq t_{\mathrm{Ae}}, \tag{19}$$

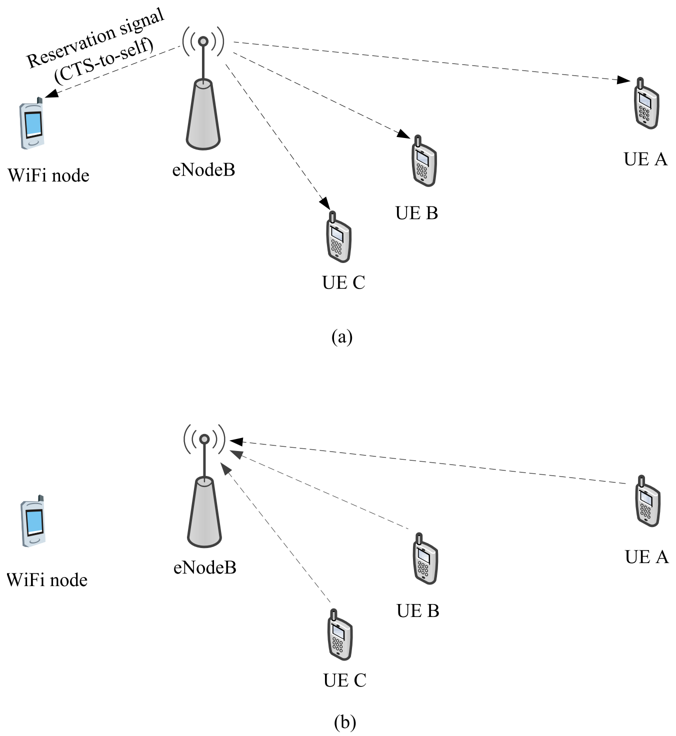


Fig. 23. Illustration of the prevention of the "hidden WiFi node problem" by the use of CTS-to-self reservation signal: (a) eNodeB broadcasts the CTS-to-self based reservation signal; (b) the hidden node (i.e., the WiFi node) remains silent per the NAV in CTS-to-self.

per the triangle inequality (see Figure 19), it follows that

$$(t_{\text{Be}} + t_{\text{CCA}}) + t_{\text{AB}} \geq (t_{\text{Ae}} + t_{\text{CCA}}), \quad (20)$$

where $t_{\text{CCA}} = 25$ μs is the CCA duration. This indicates that by the time the reservation signal from UE B reaches UE A, UE A has already finished CCA. Applying the same argument to other UEs, we come to the conclusion that the reservation signal from a UE does not preclude other UEs from accessing the resource in an FDM manner. The scheduled UEs can thus use reservation signals to secure the channel after CCA without causing blocking issues. This translates to reduced probability of disruptions by other nearby systems.

Another advantage of this uplink transmission scheme is the capability of preventing the "hidden node" problem. As depicted in Figure 23, an LAA eNodeB broadcasts a CTS-to-self reservation signal after LBT to secure the following MCOT that includes the uplink transmission duration. Therefore, the uplink reception at the eNodeB is protected from nearby WiFi transmissions *as well as* LAA systems of different operators that perform virtual sensing.

c) Hybrid ARQ: Recall that, in legacy LTE, the retransmission is triggered by the NACK signal carried on PHICH, and follows a *fixed* synchronous hybrid ARQ timeline [see Section IV-A]. In LAA, this scheme is replaced by the downlink-like asynchronous hybrid ARQ due to its flexibility that fits better to the unlicensed spectrum. PHICH is no longer used for carrying the ACK/NACK for the uplink transmission associated with an unlicensed PUSCH. In LAA, signaling is through the PDCCH/EPDCCH (licensed or unlicensed) and PUCCH of the licensed carrier.

VI. Future Research

A. Hidden Node Problem

In WiFi, the hidden node problem among WiFi nodes is addressed using the RTS and CTS signal pair. A transmitter broadcasts an RTS signal and waiting for the receiver to respond with CTS before transmission. The nearby WiFi nodes use the virtual carrier sensing mechanism to clear the channel for the period indicated in the RTS and/or CTS message. This mechanism is only used for large packets to avoid signaling overhead. In theory, LTE-U and LAA can rely on scheduling and coordination among LTE cells to avoid the hidden node problem within the same system. Hence, no specific mechanism is specified as of the current release. However, the problem of hidden node from other systems, such as WiFi systems and LTE systems from different operators still exists, although for LAA uplink this problem is less of an issue. As described in Section IV-C, the hidden WiFi node problem can be solved using the CTS-to-self based reservation signal.

In general, LTE-U/LAA systems may leverage the device-assisted detection to alleviate the hidden nodes issue among LTE-U/LAA systems and WiFi systems. A foreseeable baseline solution is the use of the UE assisted measurement via LTE RRM messages to report LTE-U traffic from different operators. The LTE-U UE that is equipped with a WiFi network listening module may also report the surrounding WiFi activities or may even take advantage of the 802.11k protocols (although not all WiFi nodes support 802.11k). Nevertheless, a complete solution to the hidden nodes issue between different systems is non-trivial and further research is necessary.

It is perhaps also worth noting here that the prevalence of severe hidden terminals in real network deployments remains unclear, especially in small cell scenarios.

B. MTC on Unlicensed Spectrum

A new narrowband MTC or M2M air interface operating on the *licensed* band has just been standardized in Release 14 LTE, coined NB-IoT, developed specifically for M2M in the *massive* IoT (mIoT) market. Due to the "massive" nature, capacity is a critical issue. The scarce licensed spectrum is ultimately the bottleneck of this technology. Undoubtedly, with the vastly-available unlicensed spectrum, unlicensed cellular technology, i.e., LTE-LAA, will play an important role in extending the LTE capacity in the M2M market. A straight forward deployment model is to use NB-IoT on licensed spectrum (e.g., the refarmed GSM spectrum) as an anchor carrier for reliable control signaling, and LAA as secondary carriers for opportunistic data traffic of large volume mIoT devices since mIoT applications are typically not sensitive to latency. Indeed, this solution leverages the large amount of free spectrum available around the 5 GHz band. Nevertheless, there are a few major challenges which may hinder the immediate application of LAA to mIoT: First is the bandwidth. As discussed in Sections II and III, there is a minimum occupied bandwidth regulatory requirement, i.e., the transmission bandwidth must be at least 80 percent of the total bandwidth (e.g., 20 MHz in 5-GHz unlicensed bands), which engenders an issue for low-cost mIoT devices whose low-cost nature prevents a device from having a wideband transceiver. Second is the coverage. Deep coverage is typically needed for mIoT and the technique is through elongated TTI. For instance, in NB-IoT, a 20-dB coverage extension over the LTE footprint is achieved via extending TTI up to 40.960 seconds on a less than 200 kHz bandwidth, whereas the maximum channel occupation time in unlicensed band is typically less than 10 ms (4 ms in Japan), which complicates coverage extension. Issues like these need to be resolved before the M2M feature can be added to the future releases of LTE-LAA.

C. D2D on Unlicensed Spectrum

The concept of D2D links as a new cell tier to wireless networks in *licensed* bands has been explored in numerous publications over the past several years [17], [19], [33], [34], [64], and has been standardized recently in 3GPP LTE. However, its practical applications are still slow. The difficulty is in part because of the D2D signaling overhead (e.g., the discovery signal transmitted periodically by UE) and potential interference to the traditional cellular connections which may degrade the service quality that is essential to the cellular service. With the LAA interface becoming available to LTE networks, it is now more natural to offload the D2D connections to unlicensed spectrum. This not only eliminates the interference to the cellular connections, but also, with the "assistance" from the licensed carrier, the management of D2D connections in unlicensed band becomes less costly and more efficient. However, technical challenges exist when it comes down to the practical implementation. Issues like how to fit the UE's discovery signals into the LAA framework on unlicensed spectrum are clearly non-trivial. More work is urgently needed.

VII. Conclusion

The wireless communication spectrum is classified as licensed and unlicensed spectrum. The limited licensed spectrum and widely-available license-exempt or unlicensed spectrum motivates the wireless industry and standard group to bring a common unified solution for both licensed and unlicensed spectrum to the ever-growing system capacity challenge. However, cellular communications on these two types of spectrum present serious technical challenges which lie mainly in the following aspects: First is the conformation to the regulatory rules of communications on the unlicensed spectrum under the cellular architecture; second is the fair sharing of medium between two very different technologies; and last is the medium access control within the cellular system, specifically the centralized multiple access among users. We review a cellular communication structure that is commonly employed in licensed spectrum, and discussed two LBT mechanisms for communications on unlicensed spectrum. We show that the load-based LBT mechanism is more robust in terms of coexistence with different types of systems, whereas the frame-based LBT is useful for certain transmissions that have strict timing requirements like the discovery signal. The baseline framework provided in this paper illustrates how a traditional cellular system can be "mutated" to operate on a different type of spectrum complying with the regulations, and most importantly, co-existing with other systems. We then provide a detailed description on such a practical system, i.e., the latest LTE-LAA technology that extends the LTE cellular technology to the unlicensed spectrum with both theoretical and practical arguments and justifications behind the design. The capability of operating on unlicensed spectrum and coexisting with other systems on the same spectrum is indeed a significant evolution of the cellular technology. There are three key components employed that provide a cellular system with such capability: First is the implementation of a load-based LBT on a cellular transmission structure. Although load-based LBT is random and asynchronous in nature, the synchronous design lends itself well to the more efficient deterministic synchronous frame-based transmission structure within the LTE network, allowing coherent interworking between unlicensed and the licensed carriers. Second is the ability to use a reservation signal, like the WiFi CTS-to-self signal that has the potential of serving seamlessly as a "common language" for interference coordination between LAA and WiFi systems, thereby truly guaranteeing no more impact on a WiFi system performance (e.g., throughput and latency) than other co-channel WiFi systems would. Last is the use of licensed carrier to provide reliable and timely signaling for centralized medium access control within the LTE-LAA system.

Acknowledgment

The authors are grateful for the excellent comments by the reviewers and the Editor that greatly improved the presentation and the contents of this paper.

References

[1] J. G. Andrews *et al.*, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[2] "Small cell enhancements for E-UTRA and E-UTRAN–physical layer aspects," 3GPP, Sophia Antipolis, France, Tech. Rep. 36.872 V12.1.0, Dec. 2013.

[3] "Study on licensed-assisted access to unlicensed spectrum," 3GPP, Sophia Antipolis, France, Tech. Rep. 36.889 V13.0.0, Jun. 2015.

[4] B. Ren *et al.*, "Cellular communications on license-exempt spectrum," *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 146–153, May 2016.

[5] A. Mukherjee *et al.*, "Licensed-assisted access LTE: Coexistence with IEEE 802.11 and the evolution toward 5G," *IEEE Commun. Mag.*, vol. 54, no. 6, pp. 50–57, Jun. 2016.

[6] L. Falconetti *et al.*, "Design and evaluation of licensed assisted access LTE in unlicensed spectrum," *IEEE Wireless Commun.*, vol. 23, no. 6, pp. 24–30, Dec. 2016.

[7] C. Cano, D. Lopez-Perez, H. Claussen, and D. J. Leith, "Using LTE in unlicensed bands: Potential benefits and coexistence issues," *IEEE Commun. Mag.*, vol. 54, no. 12, pp. 116–123, Dec. 2016.

[8] H. Cui, V. C. M. Leung, S. Li, and X. Wang, "LTE in the unlicensed band: Overview, challenges, and opportunities," *IEEE Wireless Commun.*, vol. 24, no. 4, pp. 99–105, Aug. 2017.

[9] M. G. S. Sriyananda, I. Parvez, I. Güvene, M. Bennis, and A. I. Sarwat, "Multi-armed bandit for LTE-U and WiFi coexistence in unlicensed bands," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Doha, Qatar, 2016, pp. 1–6.

[10] S. Sagari *et al.*, "Coordinated dynamic spectrum management of LTE-U and Wi-Fi networks," in *Proc. IEEE Int. Symp. Dyn. Spectr. Access Netw. (DySPAN)*, Stockholm, Sweden, 2015, pp. 209–220.

[11] M. Chen, W. Saad, and C. Yin, "Echo state networks for self-organizing resource allocation in LTE-U with uplink–downlink decoupling," *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 3–16, Jan. 2017.

[12] Y. Chen and H.-S. Oh, "A survey of measurement-based spectrum occupancy modeling for cognitive radios," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 1, pp. 848–859, 1st Quart., 2016.

[13] M. Höyhtyä *et al.*, "Spectrum occupancy measurements: A survey and use of interference maps," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 4, pp. 2386–2414, 4th Quart., 2016.

[14] X. Xing, T. Jing, W. Cheng, Y. Huo, and X. Cheng, "Spectrum prediction in cognitive radio networks," *IEEE Wireless Commun.*, vol. 20, no. 2, pp. 90–96, Apr. 2013.

[15] Y. Saleem and M. H. Rehmani, "Primary radio user activity models for cognitive radio networks: A survey," *J. Netw. Comput. Appl.*, vol. 43, pp. 1–16, Aug. 2014.

[16] R. H. Tehrani, S. Vahid, D. Triantafyllopoulou, H. Lee, and K. Moessner, "Licensed spectrum sharing schemes for mobile operators: A survey and outlook," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 4, pp. 2591–2623, 4th Quart., 2016.

[17] P. Mach, Z. Becvar, and T. Vanek, "In-band device-to-device communication in OFDMA cellular networks: A survey and challenges," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 1885–1922, 4th Quart., 2015.

[18] S. Andreev, O. Galinina, A. Pyattaev, K. Johnsson, and Y. Koucheryavy, "Analyzing assisted offloading of cellular user sessions onto D2D links in unlicensed bands," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 1, pp. 67–80, Jan. 2015.

[19] Y. Wu *et al.*, "Device-to-device meets LTE-unlicensed," *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 154–159, May 2016.

[20] A. Ali, W. Hamouda, and M. Uysal, "Next generation M2M cellular networks: Challenges and practical considerations," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 18–24, Sep. 2015.

[21] *Technical Specification Group Radio Access Network Evolved Universal Terrestrial Radio Access (E-UTRA) Physical Layer Measurements Release 14, v.14.0.0*, 3GPP Standard 36.213, Sep. 2016.

[22] *Technical Specification Group Radio Access Network Evolved Universal Terrestrial Radio Access (E-UTRA) Physical Channels and Modulation Release 14, v.14.0.0*, 3GPP Standard 36.211, Sep. 2016.

[23] D. Withers, *Radio Spectrum Management*. London, U.K.: IEE, 1999.

[24] "Limitation of radiation from industrial, scientific and medical (ISM) equipment," Int. Telecommun. Union, Geneva, Switzerland, Recommendation ITU-R SM.1056, 2007.

[25] J. Geier *et al.*, *Designing and Deploying 802.11 Wireless Networks: A Practical Guide to Implementing 802.11n and 802.11ac Wireless Networks for Enterprise-Based Applications*. Indianapolis, IN, USA: Cisco Press, 2015.

[26] *IEEE Standard for Information Technology– Local and Metropolitan Area Networks– Specific Requirements– Part 15.1a: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Wireless Personal Area Networks (WPAN)*, IEEE Standard 802.15.1, 2005.

[27] T. Temitope, *Implementation and Evaluation of a ZigBee Network*. Saarbrücken, Germany: Lambert Acad., 2014.

[28] FCC Office of Engineering and Technology. (Aug. 2015). *905462 15.401 UNII, U-NII, DFS Test Procedures*. [Online]. Available: www.apps.fcc.gov

[29] *Broadband Radio Access Networks (BRAN); 5 GHz High Performance RLAN*, Harmonized European Standard ETSI EN 301 893 v.1.7.1, 2012.

[30] J. Huang, R. A. Berry, and M. L. Honig, "Distributed interference compensation for wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 5, pp. 1074–1084, May 2006.

[31] M. K. Karakayali, G. J. Foschini, and R. A. Valenzuela, "Network coordination for spectrally efficient communications in cellular systems," *IEEE Wireless Commun.*, vol. 13, no. 4, pp. 56–61, Aug. 2006.

[32] G. Boudreau *et al.*, "Interference coordination and cancellation for 4G networks," *IEEE Commun. Mag.*, vol. 47, no. 4, pp. 74–81, Apr. 2009.

[33] K. Yang *et al.*, "Device discovery for multiHop cellular networks with its application in LTE," *IEEE Wireless Commun.*, vol. 21, no. 5, pp. 24–34, Oct. 2014.

[34] K. Zou *et al.*, "Proximity discovery for device-to-device communications over a cellular network," *IEEE Commun. Mag.*, vol. 52, no. 6, pp. 98–107, Jun. 2014.

[35] M. Wang *et al.*, "The evolution of LTE physical layer control channels," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 1336–1354, 2nd Quart., 2016.

[36] *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE Standard 802.11, 2013.

[37] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 3, pp. 535–547, Mar. 2000.

[38] "LTE-U technical report V 1.0," Alcatel-Lucent, Boulogne-Billancourt, France, Ericsson, Stockholm, Sweden, Qualcomm Technol. Inc., San Diego, CA, USA, Samsung Electron., Seongnam, South Korea, and Verizon, New York, NY, USA, Tech. Rep., Feb. 2015.

[39] *Study on License-Assisted Access Using LTE: Motivation*, document RP-140260, Ericsson, Stockholm, Sweden, Qualcomm, San Diego, CA, USA, and Huawei, Shenzhen, China, Mar. 2014.

[40] E. Almeida *et al.*, "Enabling LTE/WiFi coexistence by LTE blank subframe allocation," in *Proc. IEEE ICC*, Budapest, Hungary, 2013, pp. 5083–5088.

[41] *New Work Item on License-Assisted Access to Unlicensed Spectrum*, document RP-151045, Ericsson, Stockholm, Sweden, Huawei, Shenzhen, China, Qualcomm, San Diego, CA, USA, and Alcatel-Lucent, Boulogne-Billancourt, France, Jun. 2015.

[42] *New Work Item on Enhanced LAA for LTE*, document RP-152272, Ericsson, Stockholm, Sweden, and Huawei, Shenzhen, China, Dec. 2015.

[43] 3GPP, "Qualcomm incorporated, R1-155729: Reservation signal design for LAA," 3GPP TSG RAN WG1 meeting 82bis, Malmö, Sweden, Oct. 2015.

[44] 3GPP, "LG electronics, R1-156862: Reservation signal in LAA," 3GPP TSG RAN WG1 meeting 83, Anaheim, CA, USA, Nov. 2015.

[45] 3GPP, "NEC, R1-152689: Discussion on discontinuous transmission in LAA," 3GPP TSG RAN WG1 meeting 81, Fukuoka, Japan, May 2015.

[46] 3GPP, "Ericsson, huawei, R1-156386: RAN1 agreements and working assumptions for Rel-13 LAA," 3GPP TSG RAN WG1 meeting 82bis, Malmö, Sweden, Oct. 2015.

[47] 3GPP, "Samsung, R1-156771: Partial subframe for LAA," 3GPP TSG RAN WG1 meeting 83, Anaheim, CA, USA, Nov. 2015.

[48] 3GPP, "Alcatel-lucent Shanghai bell, alcatel-lucent, R1-157016: Support of initial partial subframe in LAA," 3GPP TSG RAN WG1 meeting 83, Anaheim, CA, USA, Nov. 2015.

[49] 3GPP, "NEC, R1-156675: Discussion on LAA DL transmission," 3GPP TSG RAN WG1 meeting 83, Anaheim, CA, USA, Nov. 2015.

[50] 3GPP, "NEC, R1-155275: Discussion on starting and ending of LAA transmission," 3GPP TSG RAN WG1 meeting 82bis, Malmö, Sweden, Oct. 2015.

[51] 3GPP, "ZTE, R1-156988: Details of design on DL frame structure and signaling for LAA," 3GPP TSG RAN WG1 meeting 83, Anaheim, CA, USA, Nov. 2015.

[52] 3GPP, "Huawei, HiSilicon, R1-155664: Discussion and comparison between partial subframe and floating subframe," 3GPP TSG RAN WG1 meeting 82bis, Malmö, Sweden, Oct. 2015.

[53] 3GPP, "LGE, samsung, NTT DOCOMO, R1-152222: WF on start and end position of DL transmission in LAA," 3GPP TSG RAN1 meeting 80bis, Belgrade, Serbia, Apr. 2015.

[54] 3GPP, "Huawei, HiSilicon, R1-155103: Candidate starting/ending positions of partial subframe and corresponding indication for LAA," 3GPP TSG RAN WG1 meeting 82bis, Malmö, Sweden, Oct. 2015.

[55] 3GPP, "CMCC, R1-156998: Further discussion on partial subframe transmission," 3GPP TSG RAN WG1 meeting 83, Anaheim, CA, USA, Nov. 2015.

[56] 3GPP, "Intel corporation, R1-156517: Further design details on partial subframe," 3GPP TSG RAN WG1 meeting 83, Anaheim, CA, USA, Nov. 2015.

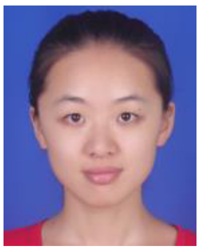
[57] 3GPP, "NEC, R1-156676: Discussion on downlink scheduling for LAA," 3GPP TSG RAN WG1 meeting 83, Anaheim, CA, USA, Nov. 2015.

[58] 3GPP, "LG electronics, R1-155390: DL synchronization and AGC issues in LAA," 3GPP TSG RAN WG1 meeting 82bis, Malmö, Sweden, Oct. 2015.

[59] 3GPP, "LG electronics, motorola, ericsson, sharp & qualcomm, R1-156245: WF on synchronization in LAA," 3GPP TSG RAN WG1 meeting 82bis, Malmö, Sweden, Oct. 2015.

[60] 3GPP, "Huawei, HiSilicon, R1-156439: LAA cell detection and synchronization for support of partial subframe," 3GPP TSG RAN WG1 meeting 83, Anaheim, CA, USA, Nov. 2015.

[61] M. Wang *et al.*, "Cellular machine-type communications: Physical challenges and solutions," *IEEE Wireless Commun.*, vol. 23, no. 2, pp. 126–135, Apr. 2016.
[62] 3GPP, "CMCC, R1-122715: Designing the scaling factor of DwPTS for new special subframe configurations," 3GPP TSG RAN WG1 meeting 69, Prague, Czech Republic, May 2012.
[63] M. Hua, M. Wang, K. W. Yang, and K. J. Zou, "Analysis of the frequency offset effect on Zadoff–Chu sequence timing performance," *IEEE Trans. Commun.*, vol. 62, no. 11, pp. 4024–4039, Nov. 2014.
[64] K. Zou *et al.*, "Discovery signal design and its application to peer-to-peer communications in OFDMA cellular networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 8, pp. 3995–4009, Aug. 2013.



Jingjing Zhang received the master's degree in electrical engineering from the Nanjing University of Science and Technology, Nanjing, China, where she is currently pursuing the Ph.D. degree, with the Wireless Networking and Mobile Communications Group, School of Electronic and Optical Engineering. She has over 10 IEEE journal publications. Her current research interests are in the areas of wireless communications and signal processing.

Mao Wang received the master's degree in biomedical engineering and the Ph.D. degree in electrical engineering and computer science from the University of Kentucky, Lexington, KY, USA. He is currently a Professor with the School of Information Science and Engineering and the National Mobile Communications Research Laboratory, Southeast University, Nanjing, China. He also is an Adjunct Professor with the School of Electronic and Optical Engineering, Nanjing University of Science and Technology, where he is the Director of the Wireless Networking and Mobile Communications Group. He holds over 90 U.S. patents and has over 30 IEEE journal publications. His research interests include communication theory and wireless networking.

Min Hua is currently pursuing the Ph.D. degree with the Wireless Networking and Mobile Communications Group, School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing, China. She has over 10 IEEE journal publications. Her current research interests are in the areas of wireless communications and signal processing.

Tingting Xia is currently pursuing the Ph.D. degree with the Wireless Networking and Mobile Communications Group, School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing, China. Her current research interests are in the areas of wireless communications and signal processing.

Wenjie Yang is currently pursuing the Ph.D. degree with the Wireless Networking and Mobile Communications Group, School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing, China. He has over 10 IEEE journal publications. His current research interests are in the areas of wireless communications and signal processing.

Xiaohu You (SM'11–F'12) received the Ph.D. degree in electrical engineering from Southeast University, Nanjing, China, in 1988. Since 1990, he has been with the National Mobile Communications Research Laboratory, Southeast University, where he is currently the Director and a Cheung Kong Scholarship Professor. He was the Premier Foundation Investigator with the China National Science Foundation. From 1999 to 2002, he was the Principal Expert of the C3G Project, responsible for organizing China's 3G mobile communications research and development activities. From 2001 to 2006, he was the Principal Expert of the National 863 Beyond 3G FUTURE Project. He has contributed over 50 IEEE journal papers and two books in the areas of signal processing, artificial neural networks, and wireless communication systems. His research interests include mobile communication systems and signal processing and its applications. He was a recipient of the China National First-Class Technical Invention Award in 2011. He is the Chairman of the IEEE Nanjing Section.