

Cognitive radio: survey on communication protocols, spectrum decision issues, and future research directions

José Marinho · Edmundo Monteiro

Published online: 7 October 2011
© Springer Science+Business Media, LLC 2011

Abstract Currently, the radio spectrum is statically allocated and divided between licensed and unlicensed frequencies. Due to this inflexible policy, some frequency bands are growing in scarcity, while large portions of the entire radio spectrum remain unused independently of time and location. Cognitive Radio is a recent network paradigm that aims a more flexible and efficient usage of the radio spectrum. Basically, it allows wireless devices to opportunistically access portions of the entire radio spectrum without causing any harmful interference to licensed users. The present document surveys the literature on Cognitive Radio. It aims to provide a comprehensive and self-contained description of this research topic area, mainly focusing on communication protocols, spectrum decision issues, and future research directions. It is a tutorial in nature and consequently does not require any previous knowledge about Cognitive Radio. Readers are only required to have some general background on wireless data networks. Emphasis is put on Cognitive Radio genesis, issues that must be addressed, related technologies, standardization efforts, the state of the art, and future research directions according to the vision of the authors.

Keywords Cognitive radio · Dynamic spectrum access · Medium access control · Spectrum decision

1 Introduction

Currently, the radio spectrum is divided into licensed and unlicensed frequencies. The licensed spectrum is for the exclusive use of designated users. For instance, it includes the UHF/VHF TV frequency bands. The unlicensed spectrum can be freely accessed by any user, following certain rules (e.g., not exceeding a defined limit for transmission power). It includes, for instance, the ISM (Industrial, Scientific and Medical) and U-NII (Unlicensed National Information Infrastructure) frequency bands. ISM is shared by technologies such as high speed wireless local area networks and cordless phones. It is used by technologies such as IEEE (Institute of Electrical and Electronics Engineers) 802.11 and IEEE 802.11 g. U-NII includes frequency bands that are used by the IEEE 802.11 a technology and by internet service providers (ISPs). Therefore, many wireless technologies operate and must coexist in the same frequency bands, and devices must compete with neighbours for the same spectrum resources. Appropriate dynamic frequency selection mechanisms have already been proposed to enable license-free wireless devices to make an efficient use of the unlicensed spectrum. However, the number of non-overlapping frequency bands in the unlicensed spectrum is limited, and increasing performance degradation cannot be avoided as it becomes more crowded, especially in densely populated areas.

Nevertheless, while the unlicensed spectrum bands are becoming more crowded, especially ISM, a report from Federal Communication Community (FCC) concluded that licensed frequency bands are often underutilized, creating

J. Marinho (✉)
ISEC—Engineering Institute of Coimbra, Polytechnic Institute of Coimbra, Coimbra, Portugal
e-mail: fafe@isec.pt

J. Marinho
CISUC—Centre for Informatics and Systems of the University of Coimbra, Coimbra, Portugal

E. Monteiro
DEI-UC—Department of Informatics Engineering,
University of Coimbra, Coimbra, Portugal
e-mail: edmundo@dei.uc.pt

temporally available spectrum opportunities that are variable in time and space [1]. In this report, measurements taken in several major cities of the United States of America (USA) prove that many portions of the spectrum below 1 GHz are not in use for significant periods of time. Other measurements, taken between January 2004 and August 2005 by the company Shared Spectrum Company (SSC), show that on the average only 5.2% of the spectrum between 30 MHz and 3 GHz is in use at six different locations in the United States of America. These measurements clearly show that large portions of the licensed spectrum remain unused by licensed users, also designated as primary or incumbent users, independently of time and space. Up to now, most of the studies concerning how different frequency bands are used have been done in the USA. However, some measurements have also been taken in other regions of the world, such as New Zealand and some European countries (e.g., the work of Lopez-Benitez et al. [2]). This can naturally be viewed as spectrum opportunities for unlicensed users, also designated as secondary users. Malicious users, which try to cause as much damage as they can, are a third type of users that can also be considered in this type of scenarios, besides primary and secondary users. According to Tsagkaris et al. [3], the underutilization of the radio spectrum is explained by the aforementioned static assignment policies, but also by what they refer as the often criticized governments' overregulation.

In this context, Cognitive Radio (CR) has emerged as one of the keys that can help addressing the aforementioned inefficient usage of the radio spectrum, without requiring the allocation of new frequency bands. It exploits unused licensed radio frequencies, often designated as spectrum holes or white spaces, opening it to secondary users. Secondary users can opportunistically use these opportunities to increase performance, without causing any harmful interference to primary users (see Fig. 1). The operating spectrum band, other transmission parameters and the access technology are dynamically and intelligently chosen by secondary users based on the spectrum availability. Spectrum mobility and spectrum handover are two main novel concepts introduced by CR, and spectrum sensing, learning based on experience and intelligent decision making algorithms are main issues in CR [4–7]. According to Akyildiz et al. [8], there are two general models for assigning spectrum usage rights in CR networks: (1) the exclusive use model, which is the classical view of opportunistic spectrum access with no harmful interference to licensed users; and (2) the commons model, which does not provide interference protection to any particular user, requiring the users to adhere to etiquettes. Beltrán et al. [9] argue that CR will also contribute to the shift to a more competitive telecommunication landscape,

where new operators will collectively and seamlessly provide customers with more flexible and dynamic spectrum arrangements. It can also be noted that the CR concept is clearly included by Chen et al. [10] in the critical path to the future wireless networks.

CR depends on the availability of SDRs (Software Defined Radios), which define a kind of radio that can be dynamically reconfigured by software. Depending on the level of sophistication of the SDR device, several parameters may be reconfigurable (e.g., the operating frequency centre, the bandwidth, the modulation scheme and the transmission power). SDRs perform a kind of operation that is often designated as Dynamic Spectrum Access (DSA), Opportunistic Spectrum Allocation (OSA), Spectrum Allocation Access or Spectrum Agile Radio. Basically, a CR is a SDR that is able to intelligently adapt its spectrum usage to the changing radio frequency environment and according to some predefined objectives (e.g., performance, availability, and reliability). Hence, efficient algorithms for learning based on experience and observation, and for decision making are highly desirable and expected. Mueck et al. [11] provide a joint discussion of SDR and CR standards. In fact, CR is highly interdisciplinary, being concerned with distinct engineering and computer science disciplines such as signal processing, communication protocols, and machine learning [12]. Therefore, CR issues can span all the layers of the communication protocol stack (see Fig. 2), but its basics are mostly limited to the physical (PHY) and medium access control (MAC) layers.

The present document aims at providing a comprehensive and self-contained review of developments in the CR research area. CR has implications on all the protocol layers (see Fig. 2), and often requires spectrum decisions to be taken according to a cross-layer and transversal approach, as illustrated by the CR engine module in Fig. 2. This survey mainly focuses communication protocol issues and the relevance of learning-based approaches for

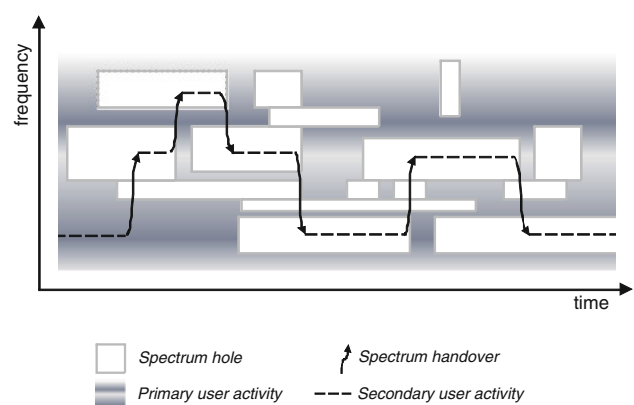


Fig. 1 Opportunistic/dynamic spectrum access

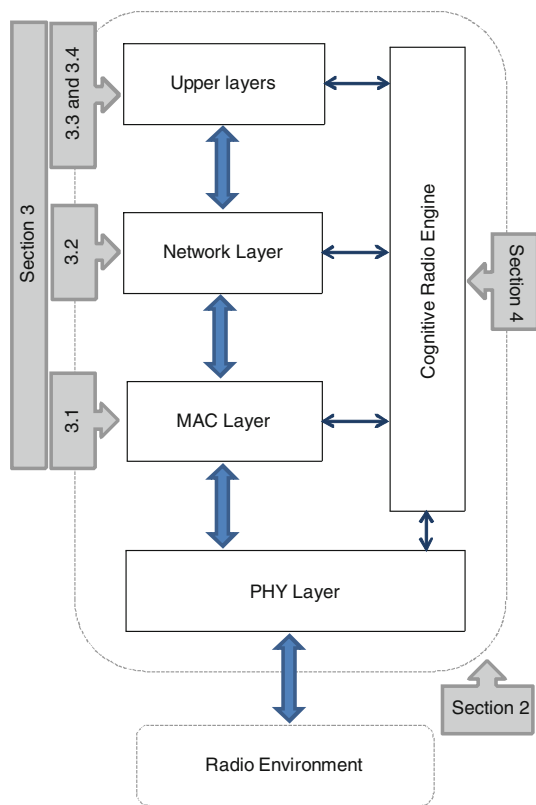


Fig. 2 Generic cognitive radio architecture

spectrum decision. It also gives some suggestions about future research directions. Hence, details about the physical layer (refer to Yucek and Arslan [13] for an initial presentation and comparison of various spectrum sensing methods for CR), and the related Software Defined Radio area remain out of the scope of this discussion. There are already a few relevant surveys in the area, such as Akyildiz et al. [4, 8], Cormio and Chowdhury [14], Wang et al. [15], and Amanna and Reed [16]. However, the current one provides a wider coverage of the area, and is more tutorial-oriented. The already published surveys are more specific and do not significantly address several issues such as learning-based approaches, and optimizations in CR MAC protocols. Therefore, the main contribution of this document is enabling anyone new to the CR area to get a global vision of it and find some interests, without going through intensive readings.

The current section has introduced the genesis and principles of CR. The next three sections are organized according to the structure illustrated by Fig. 2. In Sect. 2, a global description of the CR technology is presented. Section 3 overviews existing proposals and discussions about communication protocol issues in CR scenarios. It is concerned with the MAC sub-layer, and the upper protocol layers. Section 4 briefly discusses the relevance of learning-based approaches for spectrum decision in CR

scenarios. A few suggestions for future research directions are presented in Sect. 5. Finally, conclusions are drawn in Sect. 6.

2 Cognitive radio technology

This section provides a global description of the CR technology, focusing on architectural approaches, requirements, issues, related technologies and standardization efforts.

2.1 Objectives of CR

CR aims for an efficient utilization of the overall radio spectrum, avoiding crowded unlicensed channels while large portions of the licensed frequencies remain vacant at the same time and location. The main functionalities required for channel management in CR scenarios are spectrum sensing, spectrum decision, spectrum sharing, and spectrum mobility (see Fig. 3). This is because CR users must be able to: (1) determine the spectrum holes (spectrum sensing), possibly through coordination with other CR neighbours; (2) select the best spectrum opportunities in order to meet the user communication requirements (spectrum decision and quality of service provisioning); (3) coordinate access to the selected channels with CR neighbours (spectrum sharing); (4) switch to the selected opportunities (spectrum handover/mobility); (5) maintain seamless communication during CR operations; and (6) avoid any harmful interference to primary users.

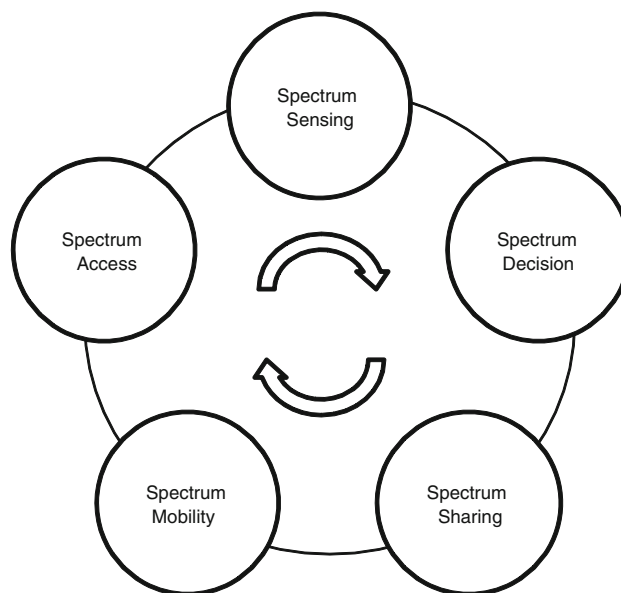


Fig. 3 Basic cognitive radio cycle

2.2 CR architectural approaches

The architecture of CR networks can either be centralized or distributed. Spectrum allocation and access are controlled by a central entity in the centralized approach (e.g., a base station in an infrastructure mode network) and by CR users in the distributed approach. Akyildiz et al. [8] designate centralized CR networks as infrastructure-based CR networks, and distributed CR networks as CR ad-hoc networks. According to Renk et al. [17], when the number of CR devices increases it becomes computationally complex for a central entity to make spectrum allocation decisions. Therefore, the distributed approach, where decision-making is processed by individual secondary users, has the ability to reduce complexity and increase network scalability.

In distributed CR networks, the secondary users, which can be fixed or mobile, need to incorporate all the CR-related capabilities, and spectrum allocation can either be achieved cooperatively or non-cooperatively. With the non-cooperative approach (device-centric), selfish users only use local policies for accessing the spectrum, do not share any interference information and try to maximize their own interests. This greedy attitude results in less communication overhead but in higher sensing inaccuracy and spectrum sharing inefficiency. In cooperative distributed CR networks, users share signalling and sensing information for spectrum allocation, serving a common goal.

2.3 Main CR issues

The last two subsections have described the main objectives of the CR technology and the two possible architectural approaches that can be applied to CR networks. This section goes further in terms of details. It describes the main issues that must be addressed by CR in order to meet its objectives.

2.3.1 Self-coexistence

One of the most important and specific issue of CR is to avoid secondary users to significantly interfere with primary users. Coexistence can be defined as the ability of a radio to coexist with other radios in the same spectrum bands using different protocols. Therefore, CR can be thought as an evolution of self-coexistence, i.e., automated coexistence based on non-manual coordination techniques. However, self-coexistence is difficult to achieve in CR scenarios because well-defined cellular architectures and frequency allocations are not provided, primary users have non-deterministic activities, and neighbouring secondary users compete for the same spectrum holes. According to

Mody et al. [18], coexistence does not require the use of cognitive techniques, but the latter can be used to facilitate the former.

Overlay and underlay are two possible spectrum access techniques. With the underlay approach, secondary users are constrained to keep transmission power below the noise floor of primary users (Interference Temperature as defined by the FCC [4]). This can be achieved through spreading the transmitted signal over a wide frequency band (e.g., higher than 500 MHz), which enables a short range high data rate to be achieved with very low transmission power [19]. UWB (Ultra Wide Band) is an example of such radio technology. The underlay approach is also designated as interference-tolerant approach and it does not experience any service interruption loss [20] or spectrum handover operation.

With the overlay approach, which requires dynamic spectrum access (DSA) and has received much more attention from the research community, a node accesses the network using a portion of the spectrum that is not being used by primary users (see Fig. 1). In this case, service interruption losses can be caused by the appearance of primary users, and can affect the performance of any layer in the communication protocol stack. For instance, Issariyakul et al. [20] study the performance issues of the transport layer in overlay CR networks. This approach requires appropriate and accurate sensing and signalling mechanisms to cope with primary user activity in CR scenarios. Sensing and signalling will be described in the next two sections. In this document, the underlay approach is not addressed and, therefore, CR specifically designates the overlay approach.

2.3.2 Spectrum sensing

Spectrum sensing aims to determine if a channel is idle or busy, in terms of primary user activity. Therefore, it is a key issue in CR. Furthermore, it can be considered a sampling process, which is related to the PHY and MAC protocol layers, and must be fast and reliable. Refer to Wellness and Mähönen [21] for some examples of theoretical formulations as well as practical approaches for spectrum sensing.

Concerning the PHY layer, Akyildiz et al. [8] briefly describe three groups of primary user detection techniques: (1) transmitter detection; (2) primary receiver detection, which is only feasible for the detection of TV receivers; and (3) interference temperature management, which is difficult to achieve. They refer that most of the current research effort focuses on the transmitter detection technique, and also describe the three schemes that can be applied to this technique, in terms of their strengths and shortcomings. These schemes are: (1) matched filter

detection; (2) energy detection; and (3) feature detection. Energy detection is the easiest scheme to implement. With it, a channel is considered busy when the strength of the detected signal level is above a certain threshold. However, energy detection requires coordinated quiet periods to avoid false alarms, as it cannot distinguish primary user activity from secondary user activity. This can introduce additional delays in the decision making process. A simple theoretical model of an energy detector is provided by Yu et al. [22].

Usually, the access to the spectrum is achieved through a “sense-before-transmit” approach (see Fig. 4), i.e., a channel is sensed before any transmission and another one is searched if it is busy. Akyildiz et al. [8] also consider that generally a periodic “sense-before-transmit” approach is used, and they designate it as the periodic sensing structure. Cormio and Chowdhury [14] discuss the trade-off between sensing and transmission times, considering that the total time is either limited or fixed (e.g., due to periodic sensing).

Cooperative MAC protocols, which enable secondary users to share sensing information with each other, are also required for an efficient and accurate detection of primary user activity. A cooperative sensing approach helps to solve the hidden primary user problem (see Fig. 5), i.e., when a channel that has been sensed available at source location experiences primary activity at the receiver side and vice versa. Therefore, it addresses problems such as the adverse effects of noise uncertainty, multi-path fading and shadowing [8, 13], which are main factors that degrade the efficiency of primary user detection in wireless networks. For instance, in the proposal of Timmers et al. [23], secondary users cannot use a channel if it is sensed busy by any of the secondary users in the network (OR-rule). With this OR-rule, SU1 and SU3 in Fig. 5 are not allowed to use channel 5 because SU2 reports it as being occupied. Malady and Silva [24] refer that the OR-rule can result in an inefficiency designated as “false spectrum access denial”, i.e., a secondary user is denied the access to a given channel despite being out of the region of potential interference to the respective primary system. They also present various clustering methods that aim to address this problem. It must be noted that the accuracy of cooperative sensing is highly dependent on the density of secondary users in the network [25].

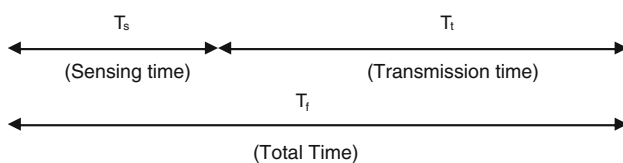


Fig. 4 The “sense-before-transmit” approach

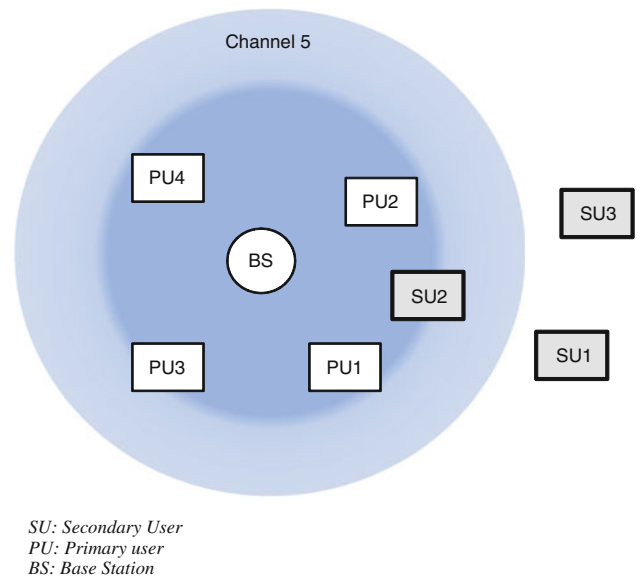


Fig. 5 The hidden primary user problem

Centralized, distributed and external sensing are possible approaches for cooperative sensing. In centralized sensing, which is appropriate for centralized architectural approaches, a central unit (e.g., a base station or an access point) collects sensing information from CR devices, processes this input for decision making through some data fusion process [25], and broadcasts the output to the CR devices and/or other central units in the network. On the other hand, in distributed sensing, which is appropriate for cooperative distributed architectural approaches, CR users share information among each other but they make their own decisions. Concerning distributed sensing, it must also be noted that the resulting signalling overhead increases with the number of secondary users. Therefore, collaborative cluster-based schemes are often considered as a means to address this issue, and any other communication overhead in wireless networks [17, 26]. In external sensing, an external agent performs the sensing and broadcasts the channel occupancy information to the CR users.

Finally, it must be referred that complicated trade-offs between trusting the detector and overlooking spectrum opportunities must exist when the access strategy considers the imperfection characteristics of spectrum detectors, i.e., missed detection and false alarm probabilities [27].

2.3.3 Signalling

It follows from the previous discussion that CR scenarios require the exchange of control information for spectrum sensing and sharing between CR devices. Most CR MAC protocols (see Sect. 3.1) use a common control channel (CCC), which facilitates signalling and neighbour

discovery in CR ad-hoc networks. CCCs (see Fig. 6) must enable the secondary users to continuously operate them without any disruption, and can be dedicated to signalling. According to Hamdaoui and Shin [28] and to Wang et al. [29], dedicating a piece of the spectrum as a “common good” (the control channel) is an absolute necessity to achieve efficient spectrum access. A CCC is easy to deploy, enables efficient broadcasting and allows distributed sensing. However, it is a potential bottleneck for network performance and scalability, as data channels remain underutilized while secondary users contend for it. There are already some proposals that aim to address this problem [14, 30, 31]. CCCs are also prone to jamming attacks by malicious users, which is one of the main classes of denial of service (DoS) attack threats in CR networks [32].

Possible candidates for CCC deployment are: (1) licensed spectrum; (2) unlicensed spectrum (e.g., ISM); and (3) UWB (see subsection 2.3.1). The unlicensed spectrum is a more feasible solution for exchanging control information, even when it becomes crowded because it is always available and control traffic has usually less performance requirements than data traffic. The licensed spectrum suffers from primary user activity, and UWB has limited coverage due to low transmission power.

2.3.4 Spectrum decision

In CR scenarios, secondary users are expected to dynamically choose the best available channels and transmission parameters [33]. However, spectrum decision is still unexplored despite being a core issue in CR [4]. In fact, CR should not be thought as being similar to generic MAC problems in existing wireless networks and limited to wireless resource allocation issues [34]. For instance, CR

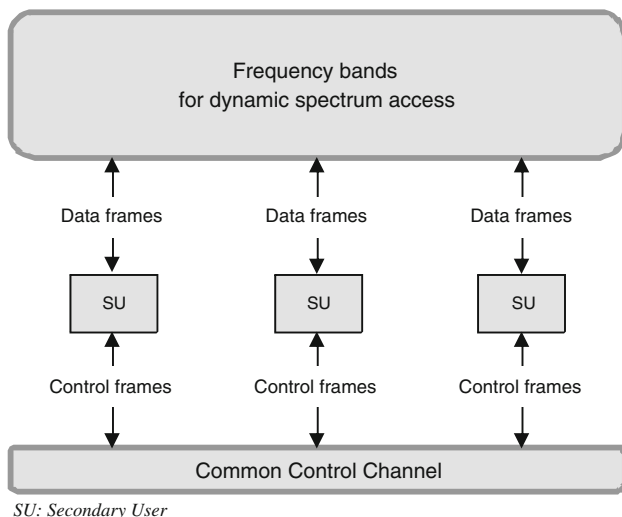


Fig. 6 Signalling based on a common control channel

MAC protocols should be able to determine the order spectrum bands must be searched (i.e., sensed) for minimizing time and energy that are needed to find a spectrum opportunity [14].

A simple yet sufficiently accurate statistical model of spectrum usage by primary users is crucial to obtain efficient solutions for tracking, selecting and accessing spectrum opportunities [19]. It essentially enables decision making even when accurate real-time information is not available (e.g. due to hardware limitations and energy constraints). Hence, collecting data about the history of the spectrum usage can help CR devices to predict future usage of the spectrum and take rewarding spectrum decisions (e.g., frequency bands to sense). The collected data must be statistically analysed and used to model the primary user activity in a given channel, such as in the work of Issariyakul et al. [20]. This is a “learning based on past experience” approach, which is further discussed in Sect. 4.

Concerning spectrum decision, it can also be referred that different frequency bands have different characteristics (e.g., path loss, link error) that should be considered at link layer for an accurate spectrum decision and to avoid spectrum outages. For instance, the area of coverage decreases as the frequency increases. Therefore, spectrum handovers to a higher frequency can result in a loss of connectivity. According to Brown and Sethi [32], available spectrum opportunities at higher frequencies can be useless due to excessively high propagation losses. There are a few proposals that consider the propagation characteristics of the frequency in usage, such as the work of Jo et al. [35].

2.3.5 Seamless spectrum handover

Seamless transition with minimum quality degradation is a fundamental goal for any spectrum handover scheme. No associated latency should be noticed by users, which may require upper layers of the protocol stack to be involved in the process. For instance, if estimation for spectrum handover latency is known in advance, an application that plays video streams from the network can buffer enough data before handover starts and deliver it while the process takes place. When QoS (Quality of Service) is a concern, observation, learning and decision making processes must be performed at the PHY, MAC, network (routing), transport (congestion control), and application layers [6]. According to Akyildiz et al. [4], there are many open research topics to be investigated concerning spectrum handover.

2.3.6 Cross-layer design

While spectrum sensing is restricted only to the PHY and MAC layers, spectrum management (e.g., spectrum

handover, decision making and scheduling) can be related to all upper layers, which makes interaction and coordination between the different layers of the protocol stack necessary. Hence, to get better performance, the strict layer-based approach is often violated in wireless networks, namely through cross-layer interactions, which is conceptually illustrated by the CR engine in Fig. 2. For instance, in the proposal of Jia and Zhang [36], routing is computed by a cross-layer entity and the results are given to the network layer that just constructs the routing table. Chen et al. [10] refer several other alternatives to the layer-based approach.

2.3.7 Energy efficiency

CR related issues must be addressed by mechanisms that are energy efficient, i.e., have limited communication and resource requirements, since most of the devices are battery powered. Hence, lightweight protocols are required for estimation, learning and decision making operations. The number of sensed channels must also be minimized through appropriate prioritisation mechanisms as sensing is one of the main sources of energy and time consumption in CR scenarios. Timmers et al. [23], Eljack et al. [37], and Grace et al. [38] are examples of works that address the energy efficiency issue.

2.4 Related technologies

As referred in previous sections, CR relies on the availability of SDRs (see Sect. 1), and MAC protocols are expected to play a central role in CR networks. However, CR MAC protocols are not based on totally new approaches. Several multichannel MAC protocols have already been proposed as extensions to wireless standards, namely IEEE 802.11-based networks, aiming to reduce interference among users and to increase throughput [39]. CR protocols essentially differ from multichannel protocols because they must be aware of the existence of non deterministic activities of higher priority users and must cope with a variable set of available channels. Hence, protocols for multichannel networks cannot be directly applied to CR networks [39].

2.5 Standards

Mody et al. [18] present a survey of CR standardization activities performed by the IEEE. IEEE 802.22 (Wireless Regional Area Network) [40] and SCC41 (Dynamic Spectrum Access Networks), formerly known as P1900, the today primary cognitive standard efforts. However, Mody et al. [18] refer other (non-CR) standards that have been addressing issues implicitly related to CR, such as the

coexistence of radios that use different protocols in the same spectrum bands. In fact, coexistence has been considered for many years within IEEE standards, initially through tedious manual coordination and frequency planning, and CR techniques can be used to facilitate coexistence [41].

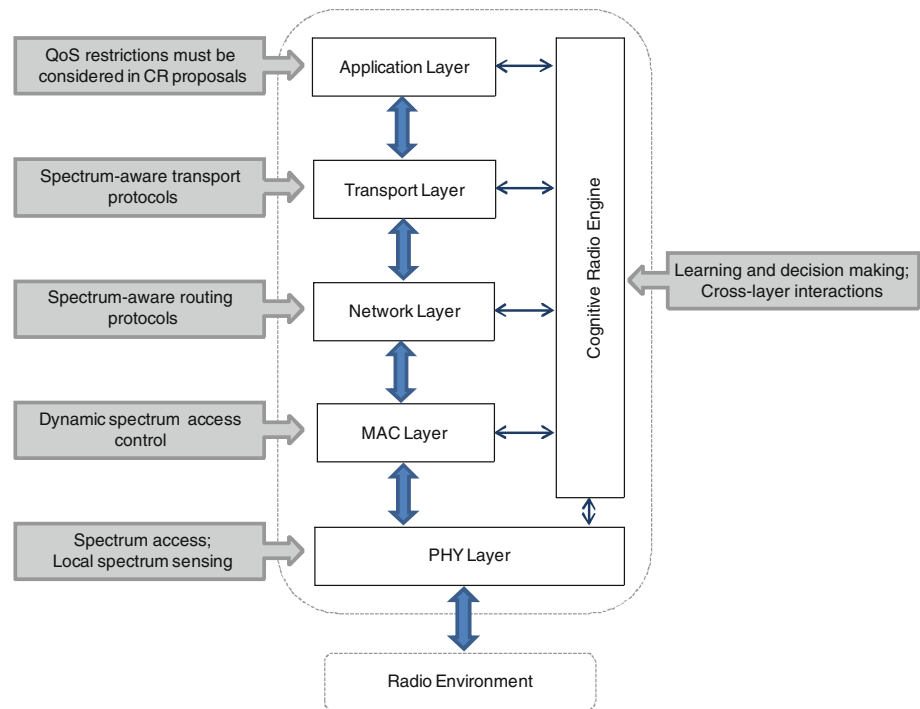
The IEEE 802.22 standard [40] defines CR techniques that are specifically targeted to enable unlicensed devices to exploit television white spaces in the VHF and UHF bands (54–862 MHz) on a non-interfering basis for the deployment of Wireless Regional Area Networks (WRAN). Its functionalities are related to both the MAC and PHY layers, and space is left for the development of new algorithms for channel sensing and classification [42]. For the protection of primary users, which is one of the main goals of channel management in CR environments, IEEE 802.22 uses both spectrum sensing and geo-location database techniques. It adopted a centralized single-hop model (each mobile device is associated with a base station), being wide coverage provided by several base stations. IEEE 802.22 also enables operations across three channels simultaneously (channel bonding) for high throughput. Refer to Ko et al. [42] for a description of the main characteristics of channel management in the IEEE 802.22 standard.

The IEEE SCC41 (Standards Coordinating Committee 41), formerly known as IEEE P1900, addresses the area of Dynamic Spectrum Access Networks (DySPAN) and aims to develop standards for next generation radio and advanced spectrum management. It includes several working groups that address specific issues such as interference and coexistence (P1900.2), and spectrum sensing (P1900.6). The first IEEE SCC41 standard was published in 27 February 2009. It provides architectural concepts and specifications for network management between incompatible wireless networks [18].

3 Communication protocols for cognitive radio scenarios

As already mentioned, the main operations in CR scenarios are mostly related to the PHY and MAC layers. Therefore, the design of appropriate MAC protocols, which are by definition responsible for controlling the access to the communication medium, is a core issue in the CR area and it has received much more attention from the research community. However, CR capabilities are also expected to have significant impacts on the performance of upper layers, and bring communication problems that can span several protocol layers [8, 43] (see Fig. 7). For instance, when QoS is a concern, observation, learning and decision making must also be performed at the network, transport

Fig. 7 Communication protocol issues in the context of Cognitive Radio scenarios



and application layers [6]. Issariyakul et al. [20] study the performance issues of the transport layer in overlay CR networks. Akyildiz et al. [8] discuss several key challenges concerning performance and open research issues faced at the network and transport layers in cognitive radio Ad-Hoc Networks (CRAHNs).

In this context, this section overviews several proposals and discussions about communication protocols in CR scenarios. The next subsections are organized according to the protocol layer-based structure.

3.1 Cognitive radio MAC protocols

Twenty two proposals (see Tables 1, 2) are used in this subsection to draw global conclusions about the main characteristics of existing CR MAC protocols. Section 3.1.1 classifies these proposals according to four fundamental parameters (see Table 1), namely the targeted network types, the adopted approaches for spectrum access, the number of required radios/transceivers, and if a common control channel is required. On the other side, Sect. 3.1.2 complements this analysis by identifying relevant supported optimizations (see Table 2).

3.1.1 Classification of existing proposals

Cormio and Chowdhury [14] investigate the characteristic features, advantages, and limiting factors of several existing CR MAC protocols, concerning sensing, time synchronization, spectrum access, and the required number of

transceivers. Both infrastructure and ad-hoc networks are considered in this work. Wang et al. [15] also provide a brief survey on a few MAC protocols for opportunistic spectrum access in CR (see Fig. 1). It can be noted that Cormio and Chowdhury [14] define three main approaches for spectrum access, i.e., for enabling multiple secondary users to determine who will access a particular available channel: (1) random access protocols, i.e., CSMA/CA (Carrier Sense Multiple Access with Collision Avoidance) like access for data and control traffic; (2) time-slotted protocols, i.e., synchronized time slots for control and data traffic; and (3) hybrid protocols, i.e., partially time-slotted and partially random access. This taxonomy is used in Table 1 from which it can be concluded that most current CR MAC proposals: (1) target ad-hoc network structures; (2) are random-based in terms of spectrum access; (3) require at least two radios/transceivers; and (4) use a common control channel (CCC) for signalling purposes (see Fig. 6). In most cases, when more than one transceiver/radio is required, one of them is dedicated for operations on the CCC.

3.1.2 Supported optimizations

The previous subsection has classified several CR MAC proposals according to four fundamental parameters. This subsection aims to complement this discussion by highlighting several relevant optimization proposals which have been identified in the analysed protocols (see Table 2).

Table 1 Classification of existing CR MAC proposals

Proposal	Network architecture	Spectrum access	Number of radios/transceivers	Control channel
Single Radio Adaptive Channel MAC (SRAC MAC) [14]	Ad-hoc	Random	1	–
Hardware Constrained-MAC (HC-MAC) [14]	Ad-hoc	Random	1	Yes
Cognitive-MAC (C-MAC) [14]	Ad-hoc	Time-slotted	Multiple	–
Partially Observable Markov Decision Process (POMDP) based MAC [14, 52]	Ad-hoc	Hybrid	1	–
Opportunistic MAC [14]	Ad-hoc	Hybrid	2	Yes
Distributed channel assignment-MAC (DCA-MAC) [14, 65]	Ad-hoc	Random	1 or 2	Yes
Distributed Frequency Agile-MAC (DFA-MAC) [5]	Mesh	Random	2	Yes
Coordinated Bandwidth Sharing-MAC (CBS-MAC) [66]	Infrastructure	Time-slotted	2	Yes
Dynamic Spectrum Allocation-MAC (DSA-MAC) [46]	Ad-hoc	Random	Multiple	Yes
Distributed Coordinated Spectrum Sharing-MAC (DCSS-MAC) [67]	Ad-hoc	Hybrid	2	Yes
Synchronized-MAC (SYN-MAC) [14, 68]	Ad-hoc	Hybrid	2	–
Opportunistic Spectrum-MAC (OS-MAC) [14, 28]	Ad-hoc	Hybrid	1	Yes
Cognitive radio CSMA/CA (C-CSMA/CA) [69]	Infrastructure	Random	1 ^a /Multiple ^b	–
Cross-Layer Based Opportunistic -MAC (CLBO-MAC) [44]	Ad-hoc	Time-slotted	2	Yes
Opportunistic Spectrum Access-MAC (OSA-MAC) [45]	–	Time-slotted	1	Yes
Statistical Channel Allocation-MAC (SCA-MAC) [47]	Ad-hoc	Random	2	Yes
Dynamic Open Spectrum Sharing MAC (DOSS MAC) [14, 31]	Ad-hoc	Random	3	Yes
Choi, Patel and Venkatesan [70]	Ad-hoc	Random	2	Yes
Ghaboosi et al. [48]	Ad-hoc	Random	2	Yes
Niyato and Hossain [6]	Mesh	Random	1	–
Jia and Zhang [36]	Ad-hoc	Random	2	Yes
Multichannel MAC protocol for CR networks (MMAC-CR) [23]	Ad-hoc	Random	2	Yes

^a Secondary users

^b Access points

Channel aggregation, which consists in combining several channels in order to create another one of greater capacity, can be a means to achieve tolerance to primary user appearance and also to enable QoS provisioning by means of throughput enhancement. Concerning QoS provisioning, it can also be concluded from Table 2 that this issue is not addressed by most proposals. The existence of backup channels is also an effective means to avoid or minimize the number of service interruption losses due to the appearance of primary users. It can be referred that several of the analysed CR MAC protocols are inspired by or based on legacy systems, especially IEEE 802.11 DCF,¹ and that some of them maintain backward compatibility. Energy efficiency is a main concern in any wireless scenario due to the existence of battery supplied devices (see Sect. 2.3). However, only two of the analysed CR MAC

¹ DCF (Distributed Coordination Function), which is based on the CSMA/CA (Carrier Sense Multiple Access with Collision Avoidance) algorithm, is one of the basic access protocols supported by the IEEE 802.11 standard. It is one of the today's main protocols for wireless local area networks.

proposals have included the energy efficiency concern in their design.

Concerning spectrum selection, it must also be noted that the majority of the proposals do not target a balanced usage of the spectrum opportunities by the secondary users (e.g., through coordination among secondary users, such as in CLBO-MAC [44]). The proposal of Niyato and Hossain [6] uses fuzzy-logic and an intelligent learning algorithm for selecting the best channel. MAC [44] and OSA-MAC [45] use random selection schemes. DSA-MAC [46], MAC [44], SCA-MAC [47], Ghaboosi et al. [48], and MMAC-CR [19] select spectrum opportunities base on simple statistics.

Cormio and Chowdhury [14] refer that further work is needed in order for CR MAC protocols to account for false alarm and missed detection probabilities, which are only addressed by HC-MAC [14] and MMAC-CR [23] in Table 2. It can also be noted that the Quality of Service (QoS) requirements of some data traffic types (e.g., multimedia streams) are not a concern for most CR MAC protocols.

Table 2 Supported optimizations

Proposal	Channel aggregation	QoS Provisioning	Backup channel	Backward compatibility	Energy efficiency	Balanced usage of opportunities	Awareness of imperfection in sensing ^b
SRAC [14]	Yes	Yes	–	–	–	–	–
HC-MAC [14]	Yes	–	–	–	–	–	Yes
C-MAC [14]	Yes	–	Yes	–	–	Yes	–
POMDP [14, 52]	–	–	–	–	–	–	–
Opportunistic MAC [14]	–	–	–	–	–	–	–
DCA-MAC [14, 65]	–	–	Yes	Yes	Yes	–	–
DFA-MAC [5]	–	–	–	Yes	–	–	–
CBS-MAC [66]	–	Yes	Yes	– ^a	–	–	–
DSA-MAC [46]	Yes	Yes	–	–	–	–	–
DCSS-MAC [67]	–	–	–	–	–	–	–
SYN-MAC [14, 68]	–	–	–	–	–	–	–
OS-MAC [14, 28]	–	–	–	–	–	Yes	–
C-CSMA/CA [69]	–	–	Yes	Yes	–	–	–
CLBO-MAC [44]	–	–	–	–	–	Yes	–
OSA-MAC [45]	–	–	–	–	–	Yes	–
SCA-MAC [47]	Yes	Yes	–	–	–	–	–
DOSS MAC [14, 31]	–	–	–	–	–	–	–
Choi et al. [70]	–	–	–	–	–	–	–
Ghaboosi et al. [48]	–	–	–	–	–	Yes	–
Niyato and Hossain [6]	–	–	–	Yes	–	–	–
Jia and Zhang [36]	–	–	–	–	–	Yes	–
MMAC-CR [23]	–	–	–	–	Yes	–	Yes

^a IEEE 802.11CSMA/CA-aware

^b False alarms and missed detections

3.2 Network layer

CR-based routing protocols must also take into account the activity of primary users and its consequences (e.g., service interruption losses) to determine the best routes. This is the main difference with traditional routing protocols. Akyildiz et al. [8] classify existing works in CR routing according to three distinct approaches: (1) routing with spectrum decision (i.e., joint selection of the spectrum and next hop); (2) routing with joint spectrum decision and primary user awareness (i.e., establishing routes that avoid locations with primary user activity); and (3) routing with joint spectrum decision and re-configurability (i.e., establishing routes that recover from primary user appearance). In their work, Wang et al. [49] also consider that there is a correlation between dynamic frequency assignment, routing and scheduling of accesses in wireless networks. Therefore, they refer that these components should be treated jointly.

Chowdhury and Felice [50] present a novel routing protocol for mobile CR ad-hoc networks which follows a

geographical forwarding approach² and is spectrum aware (i.e., aware of primary user activity). These authors also describe some related works about centralized and distributed routing protocols for CR networks, as well as geographic routing protocols. Concerning existing distributed routing protocols for CR, they refer that they scale well, but have a relevant problem: they first start identifying the best paths, and only then choose the preferred channels along them. These authors also argue that path and channel decision should be made together.

When a primary user is detected, at either end of a link, the routing protocol must take the decision of either switching to another channel in the affected link or determining an alternate path. This decision must take into account the end-to-end delay or any other end-to-end

² In geographic routing, a node knows the location of the destination and that of the candidate forwarding nodes within its range, and therefore can choose the next hop that is closer towards the destination.

performance metric, i.e., if the cost of the detour is less than the added and temporally delay in switching to a new channel [8, 50]. When disconnection is not caused by a primary user appearance (e.g., due to node mobility), the only alternative is determining another path.

Finally, concerning routing in CR networks, the work of Guan et al. [43], which is based on a different approach, can be referred. Instead of designing a new routing algorithm, these authors propose a middle-ware scheme (located between the CR function and routing) that controls the topology in CR Mobile ad-hoc Networks (CR-MANETs) based on link duration prediction. This approach is designated as PCTC (Prediction-based Cognitive Topology Control), and enables taking advantage of existing routing algorithms for MANETs. Basically, PCTC constructs efficient and reliable topologies in order to reduce re-routings and improve end-to-end performance.

3.3 Transport layer

In CR networks, a secondary user is not able to forward packets during sensing. Therefore, sensing periods must also be considered at the transport layer in order to avoid excessive retransmissions and packet losses on the paths with any node in sensing state, especially for multi-hop distributed networks. This makes the interaction between MAC and transport entities necessary. Stopping transmission at the transport layer or reducing its rate towards an optimal value, which avoids buffer overflow at intermediate nodes while maintaining transmission, are two possible approaches. Globally, the transport protocols need to be spectrum aware in CR scenarios and, therefore, require new algorithms (e.g., for congestion window scaling in TCP) [8]. Akyildiz et al. [8] describe a TCP-based protocol for CRAHNs, which they refer to be the first work aiming to address the transport layer challenges in CRAHNs. Finally, the work of Issariyakul et al. [20] can also be referred as it addresses the performance issues of the transport layer in CR networks.

3.4 Application layer

According to Yu et al. [51], the perceived reduction of the QoS at the application layer by secondary users can limit the success of CR technologies. These authors refer that most of the previous work about CR considers maximizing the throughput of secondary users, and mostly ignore other QoS measures (e.g. distortion for multimedia applications). They argue that recent work in cross-layer design shows that maximizing throughput does not necessarily benefit QoS at the application layer for some applications, such as video. Therefore, they consider that QoS at the application layer must be carefully considered in CR proposals. They

also describe in their work a design approach, as an extension to POMDP based MAC [52] (see Sects. 3.1, 4.3), that jointly optimize application layer QoS for multimedia transmission together with spectrum access and spectrum sensing in CR networks.

4 Spectrum decision in CR

In distributed CR networks, a CR device can be defined as a software defined radio (see Sect. 1) empowered with an independent CR engine (see Figs. 2, 7). The CR engine is essentially responsible for learning and spectrum decisions, and it is conceptually composed of a knowledge base, a reasoning engine and a learning engine [53]. At any time, conclusions are generated based on the information that is defined in the knowledge base. The learning engine is responsible for manipulating the knowledge, based on experience, and its outcomes are intended to be accessed in the future by the reasoning engine for decision making. At given times, the reasoning engine looks at the current state and determines which actions are executable under that state. The action which maximizes an objective function is executed, and the output it produces is used to update the knowledge base.

After this brief conceptual description of the CR engine, the next subsection highlights the expected benefits of learning based on past experience and prediction in CR scenarios, concerning spectrum decision. Then, modelling of primary user traffic is discussed. Finally, the main characteristics of some CR proposals with prediction-based capabilities are described.

4.1 Learning based on past experience and prediction in CR scenarios

In CR scenarios, secondary users should be able to dynamically, intelligently and on the fly decide on their operating frequency bands and other transmission parameters based on: (1) the observed spectrum availability/radio environment; (2) policies; (3) capabilities; (4) predefined goals (e.g., performance and QoS requirements); and (5) learning based on past experience. The relevance of learning based on past experience can be easily justified if we consider its resulting benefits in the context of the “sense-before-transmit” process (see Sect. 2.3.2). Basically, it enables to model and, consequently, predict (i.e., probabilistically determine) busy and idle times in terms of primary user activity. Therefore, some of the consequences of predictive spectrum decision are: (1) a decrease of time and energy spent to find an idle channel before any transmission, because channels can be prioritized according to their probabilities of availability; (2) a decrease in the

number of spectrum handovers and service interruption losses, because channels can be prioritized according to their expected durations of availability; and (3) a decrease in terms of interference to primary users, because primary user appearance can be anticipated during transmission. However, other less obvious benefits can result from traffic prediction in CR scenarios. For instance, according to Guan et al. [43], routing should be forward-looking (i.e., proactive). Chowdhury and Felice [50], which present a novel routing protocol for mobile CR ad-hoc networks, also consider that their work can be further enhanced by incorporating a learning based approach that determines the characteristics of primary user activity in every link's region.

According to Clancy et al. [53], the application of machine learning to CR Networks can be defined as follows: "CR users should be able to remember lessons learned from the past and act quickly in the future". They also refer that, usually, the most intelligence is put behind the learning and decision making processes, the slower they are and the higher the consumption of resources. Therefore, it is important to select the techniques taking into account the constraints that are inherent to CR scenarios (e.g., real-time operations and limited capacity in terms of processing, memory and energy).

The next subsection discusses primary user traffic modelling, a key issue in the context of learning based on past experience, and prediction in CR scenarios, the issues which have been discussed in this subsection.

4.2 Primary user traffic modelling

According to Akyildiz et al. [8], models about primary user activity must be incorporated in the characterization of the channels. This essentially enables prediction, probabilistic reasoning and intelligent spectrum decisions. Wang et al. [25] refer that several research works have indicated that channel occupancy exhibits behavioural patterns and can be statistically modelled. However, different traffic models apply to different applications (e.g., voice communication, video data, and general packet data) [54]. In CR scenarios, primary user traffic is usually modelled as a binary sequence of alternating busy (ON) and idle (OFF) states with variable durations. Therefore, the main concern in CR scenarios is predicting the durations of the present and future ON and OFF periods.

There are two basic classes of traffic in wireless scenarios: (1) deterministic patterns; and (2) stochastic (i.e., random or non-deterministic) patterns. Stochastic traffic can only be described in statistical terms (e.g., exponentially distributed) and usually vary slowly in data networks [55]. Deterministic traffic includes fixed or periodic OFF and ON times. According to Höyhty et al. [56], different

types of traffic patterns require different prediction methods/algorithms. Therefore, the type of traffic pattern must be determined during an initial learning phase (discovery) through the gathering of enough data. According to Akyildiz et al. [8], most of the research in CR assumes that the lengths of ON and OFF states are exponentially distributed, such as in the works of Issariyakul et al. [20], and Wang et al. [25]. However, Wang et al. [25] also refer that this assumption may not be valid in real world situations. Akyildiz et al. [8] briefly describe the main characteristics of some proposals about primary user activity modelling and refer that some of them are not practical. Amanna and Reed [16] also refer that some existing models are so complex that they could never be practical.

Finally, concerning traffic modelling for CR scenarios, time series-based approaches can also be referred. For instance, according to Wang and Salous [57], a sequence of spectrum measurements, typically performed at regular intervals, can be analysed by time series-based approaches in order to enable prediction.

4.3 Cognitive radio proposals with prediction-based capabilities

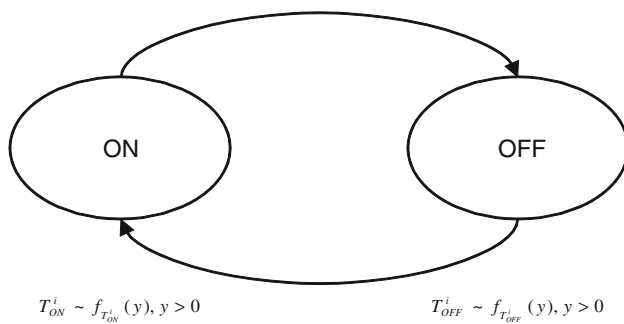
Despite the aforementioned relevance of prediction for spectrum decision, existing CR proposals mostly perform channel selection based on random and simple statistics-based approaches (see Sect. 3). According to Höyhty et al. [56], prediction based on traffic pattern characterization has not been explored much in the literature. However, there are some machine learning-based proposals in the CR area that aim to exploit the characteristics of primary traffic, and estimate the duration of occupancy times by primary users, through a statistical analysis of the gathered data [55]. The aim is to evaluate, based on experience, the probability of availability of a frequency band within a specific time period. Table 3 summarizes the main features of some of the few CR proposals that include prediction-based capabilities, in terms of their main objectives, global approaches, and base estimation methods.

In the work of Höyhty et al. [56], as the traffic of a channel can change over time, estimations only consider data that has been gathered within a limited (sliding) time-window. This mechanism aims to limit the history of events that are remembered at any given time in order to adapt to changes in spectrum occupancy. A similar approach is followed by Tsagkaris et al. [3]. In the work of Jiang et al. [58], a higher value for a weight means a more successful access to the corresponding channel in the past. These authors also analyse the performance of the three different weighting schemes [59].

Kim and Shin [60] define an estimation method which is used to capture time periods in which secondary users can

Table 3 CR proposals with prediction-based capabilities

Proposal	Main objective	Global approach	Estimation method
Höyhtyä et al. [56]	Minimizing the interference to primary users	Prediction of the future idle times and switching to better channels before the appearance of primary users on the current channel	Pattern type detection and availability time prediction
Jiang et al. [58]	Reducing the number of spectrum sensing operations	Every user assigns to every channel a weight which is updated every time the channel is accessed, through the assignment of a reward (increase) or a punishment (decrease), depending on the success of the action	Decentralized reinforcement learning scheme
Kim and Shin [60]	Efficient discovery of spectrum opportunities	Characterization of the usage of a channel by its primary users	Semi-Markov process with two states
Tsagkaris et al. [3]	Increasing the achievable data rate	Learning schemes that behave as predictors for the data rate that can be achieved when specific configurations are applied to the radio	Artificial neural networks
POMDP based MAC [52]	Optimized selection of the next channels to be sensed/accessed	Past history is accumulated and used in order to learn which channel is best suited for long term use	Partially observable Markov decision process
DOSP [25]	Improving the overall likelihood of channel access	Long-term statistical information is exchanged amongst cooperative secondary users	Statistics-based

**Fig. 8** Semi-Markov process

transmit without causing any harmful interference to primary users (i.e., OFF times). The duration of busy times T_{ON}^i and idle times T_{OFF}^i of a given channel i are modelled through the probability density functions (see Fig. 8). Their approach is purely based on statistics, and they consider that the channel parameters for the distribution of ON and OFF periods vary slowly over time. This enables secondary users to track their variations through sampling and estimation making. The ANN which is proposed by Tsagkaris et al. [3] is used to implement the learning engine of the CR engine (see Fig. 2). In the partially observable Markov decision process (POMDP) based MAC protocol proposal, which is also briefly described by Cormio and Chowdhury [14], Zhao et al. [52] consider that the network state cannot be fully observed due to partial spectrum sensing or sensing error.

Finally, Wang et al. [25] present the DOSP (Dynamically Optimized Spatiotemporal Prioritization) algorithm which is statistics-driven. The statistical information which

is exchanged between the secondary users can be fused with local channel availability statistics, through what the authors designate as dynamic statistical learning process. The resulting information is then used to prioritize the channels. The DOSP algorithm uses four types of channel occupancy statistics, as well as knowledge about the success of past decisions.

5 Future research directions

Some suggestions will now be drawn about future research directions that must be pursued in order to turn CR into a mature area. As referred previously, CR operations are mostly related to the MAC sub-layer. Therefore, some research directions are first suggested in the context of this protocol entity. However, CR also requires the development of optimized solutions for spectrum decision and mobility, which can relate to any protocol layer and, therefore, is conceptually connected to the CR engine (see Fig. 2). Thus, this issue is also discussed. Finally, as it is fundamental do accurately validate the current and future CR proposals, the last subsection is devoted to this topic.

5.1 MAC sub-layer

This subsection suggests some future research directions that are mostly concerned with the MAC sub-layer in CR scenarios.

5.1.1 Accurate sensing

Sensing (see Sect. 2.3) is the primary functionality that is required in CR scenarios. The other main CR operations, such as traffic modelling, traffic prediction, protection of primary systems, and spectrum decision, rely on it. Currently, accurate sensing (i.e., sensing without missed detections and false alarms), which is not trivial to achieve in CR scenarios, remains an open issue. New sensing solutions, which consider the trade-off between primary user protection, signalling overhead, and spectrum efficiency, must be proposed. For instance, according to Mueck et al. [11], current directions for spectrum sensing include localization techniques in order to generate a radio environment map for the presence and activity of primary users. It can also be noted that Renk et al. [17] refer that clever clustering techniques are needed as they help to reduce the search period enormously.

5.1.2 Alternatives to the common channel

It has been outlined in previous sections that most of the existing CR proposals are based on a common control channel that is used for coordination and efficient spectrum sensing. However, this approach can saturate and is prone to jamming attacks. Therefore, more efficient and robust alternatives must be investigated. Cormio and Chowdhury [30] refer several related works that aim to set up and maintain reliable CCCs, while addressing several inherent challenges. They also highlight several drawbacks concerning these works. AMRCC (Adaptive Multiple Rendezvous Control Channel), a CCC design for CR ad-hoc networks based on frequency hopping, which is proposed by Cormio and Chowdhury [30], aims to improve the network performance by overcoming these issues.

5.1.3 CR in wireless mesh and Ad-hoc networks

Wireless Mesh (WMN) and Ad-hoc Networks are expected to take advantage of new technologies such as distributed CR through better and flexible spectrum utilization and dynamic variable frequency planning. This evolution has a strong impact on MAC and routing protocols designing [61]. Hence, CR can be considered a key enabler for leveraging the frequency planning effort, and addressing other specific issues.

5.1.4 CR in emergent network types

The potential benefits of CR in the context of emergent wireless networks must also be investigated. For instance, opportunistic networks are an emerging area that defines

networks where connections can be very dynamic and are based on sporadic and intermittent contacts. It differs from legacy mobile ad-hoc network (MANET) where it is assumed that an end-to-end path always exists between the source and the destination. Opportunistic networks addresses a scenario based on moving devices that continuously change their connections based on sensed opportunities and where hard (perceived) handovers are expected to occur. CR techniques can possibly be applied in this context to help reducing the number of perceived communication service interruptions and increase performance.

5.2 Spectrum mobility and decision

This subsection suggests some research directions related to issues that can potentially span all the protocol layers, and are conceptually located in the CR engine (see Fig. 2).

5.2.1 Seamless spectrum mobility

Seamless spectrum mobility is a major requirement in CR scenarios, as any lack of QoS assurance for secondary users in CR networks is undesirable, especially for some classes of data traffic (e.g., multimedia streams and real-time traffic). It can result in unacceptable service interruptions or degradation of performance. Spectrum handovers also affect link state parameters and, therefore, can be wrongly perceived as network instability (e.g., congestion or link errors) by protocols in upper layers (e.g., transport and routing protocols). According to Akyildiz et al. [4], there are many open research topics to be investigated concerning spectrum handovers, and Sherman et al. [41] state that the type of the information to be accessed, QoS and security requirements for data streams should be considered. For instance, to achieve seamless spectrum mobility, an approach can consist in buffering enough data to be played during the spectrum handover process [4]. This in turn brings new challenges, such as estimation of spectrum handover duration, how to anticipate data gathering, and at which protocol layer should buffering be performed.

5.2.2 Proactive spectrum decision

Intelligent decision making mechanisms are expected in order to be possible to deliver the maximum performance to secondary users and avoid harmful interference to primary users. In this context, the applicability and feasibility of learning based on past experience and observation are often considered a core issue in CR. Several proposals concerning these issues do exist (see Sect. 4), but most of them were not integrated with CR MAC protocols. Therefore, their practicality was not properly evaluated.

In Sect. 4, POMDP [52] and the work of Jiang et al. [58] are two exceptions. Evaluating and validating the traffic modelling mechanisms using real measurements, which is neither a common practice, is also another important issue that must be addressed in order to draw more convincing conclusions about prediction in CR scenarios. Finally, Amanna and Reed [16] can also be referred in this discussion as they consider that some of the open research issues in CR are: (1) developing more proactive approaches that seek to make changes before they are needed; (2) adapting the use of prediction algorithms used in other research areas to the CR area; and (3) investigating the application of lesser known Artificial Intelligence algorithms.

5.3 Validation of CR proposals

Most CR proposals referred in Sect. 3.1 were validated through simulation or analytically, which is not totally convincing in many situations. Chowdhury and Melodia [12] argue that current simulators make several simplifying assumptions concerning the characteristics of real systems, resulting in questionable results. Newman et al. [62] also refer that CR research has been mostly limited to theoretical frameworks and simulations, or in few cases to prototypes for small-scale experiments. Theoretical models and simulations enable to achieve a view of what the limits could be, i.e., the performance goals. However, experimental testbeds enable an evaluation and refinement of the models. They provide much more practical results, especially concerning practicality, limitations, and the proof that harmful interference to primary users is avoided [62]. Hence, the existence of experimental platforms and testbeds is highly desirable, and a few proposals already exist, such as the framework presented by Jia and Zhang [36], Iris [63], ORBIT (Open Access Research Testbed for Next-Generation Wireless Networks) [64], and CORNET (CR Network) [62].

6 Conclusions

The CR area aims to enable an efficient utilization of the radio spectrum and is still in its infancy. In this survey, it has been comprehensively described under different perspectives, addressing layer-2 issues, learning based on past experience, and implications on upper layer protocols. The main objective of this document was to provide the readers a global vision of CR concerning its principles, present state and possible future evolution. It was shown that several challenging issues have already been satisfactorily addressed and that many still need further investigation, making CR an open and interdisciplinary research area.

According to our vision of CR, which was specified in the previous sections, some of the main open challenges for the deployment of efficient CR scenarios are: (1) accurate sensing; (2) proactive spectrum decision; (3) efficient and balanced usage of spectrum holes; (4) energy efficiency; (5) robust and scalable alternatives to the common control channel approach; (6) seamless spectrum mobility; and (7) QoS provisioning. Integrating practical learning capabilities with CR proposals is also a core challenge in CR, as this enables making decisions when accurate real-time information is not available, i.e., through prediction, and can result in several benefits that were outlined in Sect. 4. All these issues require appropriate mechanisms that are intimately linked but not limited to MAC protocols. These mechanisms must take into account complexity, processing rates (e.g., CR requires real-time operations) and resource requirements (e.g., energy consumption). Besides the issues that have been discussed with some detail throughout the previous sections, other domains must also be considered, such as business models to support CR and regulation issues.

It can be concluded that CR is without any doubt in the critical path to the wireless networks of the future. However, a significant amount of work remains to be done.

Acknowledgments The authors would like to thank the anonymous reviewers for their constructive comments and suggestions which have significantly contributed to improve this paper. The first author is also grateful to the support which was provided in the context of PROTEC, a Portuguese PhD supporting program.

References

1. FCC Spectrum Policy Task Force. (2002). Report of the spectrum efficiency working group.
2. Lopez-Benitez, M., Umberto, A., & Casadevall, F. (2009). Evaluation of spectrum occupancy in Spain for cognitive radio applications. In *69th Conference on vehicular technology, 2009* (pp. 1–5). VTC Spring 2009, IEEE.
3. Tsagkaris, K., Katidiotis, A., & Demestichas, P. (2008). Neural network-based learning schemes for cognitive radio systems. *Computer Communications (Elsevier)*, *31*(14), 3394–3404.
4. Akyildiz, I., Lee, W., Vuran, M., & Mohanty, S. (2008). A survey on spectrum management in cognitive radio networks. *Communications Magazine, IEEE*, *46*, 40–48.
5. Ghaboosi, K., Latva-aho, M., & Xiao, Y. (2008). A distributed multi-channel cognitive MAC protocol for IEEE 802.11 s wireless mesh networks. In *3rd International conference on cognitive radio oriented wireless networks and communications, 2008* (pp. 1–8). CrownCom 2008.
6. Niyato, D., & Hossain, E. (2009). Cognitive radio for next-generation wireless networks: an approach to opportunistic channel selection in IEEE 802.11-based wireless mesh—[accepted from open call]. *Wireless Communications, IEEE*, *16*, 46–54.
7. Vuran, M., & Akyildiz, I. (2007). A-MAC: Adaptive medium access control for next generation wireless terminals. *IEEE/ACM Transactions Network*, *15*, 574–587.

8. Akyildiz, I., Lee, W., & Chowdhury, K. (2009). CRAHNs: Cognitive radio ad hoc networks. *Ad Hoc Networks (Elsevier)*, 7(5), 810–836.
9. Beltrán, F., Gutiérrez, J., & Melús, J. (2010). *Technology and market conditions towards a new competitive landscape in the wireless access market*. IEEE Communications Magazine, pp. 46–52.
10. Chen, T., Zhang, H., Matinmikko, M., & Katz, M. (2008). *CogMesh: Cognitive wireless mesh networks* (pp. 1–6). GLOBECOM workshops, 2008 IEEE.
11. Mueck, M., et al. (2010). ETSI reconfigurable radio systems: status and future directions on software defined radio and cognitive radio standards. *IEEE Communications Magazine*, 48(9), 78–86.
12. Chowdhury, K., & Melodia, T. (2010). Platforms and test beds for experimental evaluation of cognitive ad hoc networks. *IEEE Communications Magazine*, 48(9), 96–104.
13. Yucek, T., & Arslan, H. (2009). A survey of spectrum sensing algorithms for cognitive radio applications. *Communications Surveys & Tutorials, IEEE*, 11, 116–130.
14. Cormio, C., & Chowdhury, K. (2009). A survey on MAC protocols for cognitive radio networks. *Ad Hoc Networks (Elsevier)*, 7(7), 1315–1329.
15. Wang, H., Qin, H., & Zhu, L. (2008). A survey on MAC protocols for opportunistic spectrum access in cognitive radio networks. In *International conference on computer science and software engineering* (pp. 214–218).
16. Amanna, A., & Reed, J. (2010). Survey of cognitive radio architectures. In *Proceedings of the IEEE southeast conference 2010* (pp. 292–297).
17. Renk, T., Kloeck, C., Burgkhardt, D., Jondral, F., Grandblaise, D., Gault, S., et al. (2008). Bio-inspired algorithms for dynamic resource allocation in cognitive wireless networks. *Mobile Networks and Applications*, 13(5), 431–441.
18. Mody, A., Sherman, M., Martinez, R., Reddy, R., & Kiernan, T. (2008). Survey of IEEE standards supporting cognitive radio and dynamic spectrum access. In *Military communications conference, 2008* (pp. 1–7). MILCOM 2008. IEEE.
19. Zhao, Q., & Swami, A. (2007). A survey of dynamic spectrum access: Signal processing and networking perspectives. In *IEEE international conference on acoustics, speech and signal processing, 2007* (pp. IV-1349–IV-1352). ICASSP 2007.
20. Issariyakul, T., Pillutla, L., & Krishnamurthy, V. (2009). *Tuning radio resource in an overlay cognitive radio network for TCP: Greed Isn't good* (pp. 57–63). IEEE Communications Magazine.
21. Wellens, M., & Mähönen, P. (2010). Lessons learned from an extensive spectrum occupancy measurement campaign and a stochastic duty cycle model. *Mobile Networks and Applications*, 15(3), 461–474.
22. Yu, F., Tang, H., Huang, M., Li, Z., & Mason, P. (2009). Defense against spectrum sensing data falsification attacks in mobile ad hoc networks with cognitive radios. In *Military communications conference* (pp. 1–7). MILCOM 2009. IEEE.
23. Timmers, M., Pollin, S., Dejonghe, A., Van der Perre, L., & Catthoor, F. (2010). A distributed multichannel MAC protocol for multihop cognitive radio networks. *IEEE Transactions on Vehicular Technology*, 59, 446–459.
24. Malady, A., & da Silva, C. (2008). Clustering methods for distributed spectrum sensing in cognitive radio systems. In *Military communications conference, 2008* (pp. 1–5). MILCOM 2008, IEEE.
25. Wang, X., Wong, A., & Ho, P. (2010). Dynamically optimized spatiotemporal prioritization for spectrum sensing in cooperative cognitive radio. *Wireless Networks*, 16(4), 889–901.
26. Baddour, K., Üreten, O., & Willink, T. (2010). A distributed message-passing approach for clustering cognitive radio networks. *Wireless Networks*. doi:10.1007/s11277-010-0010-z.
27. Zhao, Q., & Sadler, B. (2007). A survey of dynamic spectrum access. *Signal Processing Magazine, IEEE*, 24, 79–89.
28. Hamdaoui, B., & Shin, K. (2008). OS-MAC: An efficient mac protocol for spectrum-agile wireless networks. *IEEE Transactions on Mobile Computing*, 7, 915–930.
29. Wang, J., Abolhasan, M., Safaei, F., & Franklin, D. (2007). A survey on control separation techniques in multi-radio multi-channel MAC protocols. In *International Symposium on communications and information technologies, 2007* (pp. 854–859). ISCIT '07.
30. Cormio, C., & Chowdhury, K. (2010). Common control channel design for cognitive radio wireless ad hoc networks using adaptive frequency hopping. *Ad Hoc Networks (Elsevier)*, 8(4), 430–438.
31. Ma, L., Han, X., & Shen, C. (2005). Dynamic open spectrum sharing MAC protocol for wireless ad hoc networks. In *First IEEE International Symposium on new frontiers in dynamic spectrum access networks, 2005* (pp. 203–213). DySPAN 2005.
32. Brown, T., & Sethi, A. (2008). Potential cognitive radio denial-of-service vulnerabilities and protection countermeasures: A multi-dimensional analysis and assessment. *Mobile Networks and Applications*, 13(5), 516–532.
33. Sharma, A., & Belding, E. (2008). FreeMAC: Framework for multi-channel mac development on 802.11 hardware. In *Proceedings of the ACM workshop on programmable routers for extensible services of tomorrow* (pp. 69–74). Seattle, WA: ACM.
34. Ji, Z., & Liu, K. (2007). Cognitive radios for dynamic spectrum access—dynamic spectrum sharing: A game theoretical overview. *Communications Magazine, IEEE*, 45, 88–94.
35. Jo, O., Choi, H., & Cho, D. (2009). Seamless spectrum handover improving cell outage in cognitive radio systems. In *4th International conference on cognitive radio oriented wireless networks and communications, 2009* (pp. 1–6). CROWNCOM '09.
36. Jia, J., & Zhang, Q. (2009). A testbed development framework for cognitive radio networks. In *IEEE international conference on communications, 2009* (pp. 1–5). ICC '09.
37. Eljack, S., Huang, B., Tu, L., & Zhang, P. (2009). Synchronized multi-channel cognitive MAC protocol with efficient solutions for second spectrum access. In *Ubiquitous, Symposia and workshops on autonomic and trusted computing, 2009* (pp. 477–481). UIC-ATC '09.
38. Grace, D., Chen, J., Jiang, T., & Mitchell, P. (2009). Using cognitive radio to deliver 'Green' communications. In *4th International conference on cognitive radio oriented wireless networks and communications, 2009* (pp. 1–6). CROWNCOM '09.
39. Yau, A., Komisarczuk, P., & Teal, P. (2008). On multi-channel MAC protocols in cognitive radio networks. In *Telecommunication networks and applications conference, 2008* (pp. 300–305). ATNAC 2008, Australasian.
40. IEEE 802.22 WRAN WG Website, (Online). (2009). Available: <http://www.ieee802.org/22/>. Accessed 2 Dec 2009.
41. Sherman, M., Mody, A., Martinez, R., Rodriguez, C., & Reddy, R. (2008). IEEE Standards supporting cognitive radio and networks, dynamic spectrum access, and coexistence. *Communications Magazine, IEEE*, 46, 72–79.
42. Ko, G., Franklin, A., You, S., Pak, J., Song, M., & Kim, C. (2010). Channel management in IEEE 802.22 WRAN systems. *IEEE Communications Magazine*, 48(9), 88–94.
43. Guan, Q., Yu, F., Jiang, S., & Wei, G. (2010). Prediction-based topology control and routing in cognitive radio mobile Ad Hoc

- networks. *IEEE Transactions on Vehicular Technology*. doi: [10.1109/TVT.2010.2069105](https://doi.org/10.1109/TVT.2010.2069105).
44. Su, H., & Zhang, X. (2008). Cross-layer based opportunistic MAC protocols for QoS provisioning over cognitive radio wireless networks. *IEEE Journal on Selected Areas in Communications*, 26, 118–129.
 45. Le, L., & Hossain, E. (2008). A MAC protocol for opportunistic spectrum access in cognitive radio networks. In *Wireless communications and networking conference, 2008* (pp. 1426–1430). WCNC 2008, IEEE.
 46. Joe, I., & Son, S. (2008). Dynamic spectrum allocation MAC protocol based on cognitive radio for QoS support. In *Japan-China joint workshop on frontier of computer science and technology, 2008* (pp. 24–29). FCST '08.
 47. Hsu, A., Weit, D., & Kuo, C. (2007). A cognitive MAC protocol using statistical channel allocation for wireless Ad-Hoc networks. In *Wireless communications and networking conference, 2007* (pp. 105–110). WCNC 2007, IEEE.
 48. Ghaboosi, K., MacKenzie, A., DaSilva, L., Abdallah, A., & Latva-Aho, M. (2009). A channel selection mechanism based on incumbent appearance expectation for cognitive networks. In *Wireless communications and networking conference, 2009* (pp. 1–6). WCNC 2009, IEEE.
 49. Wang, X., & Garcia-Luna-Aceves, J. (2010). Collaborative routing, scheduling and frequency assignment for wireless Ad Hoc networks using spectrum-agile radios. *Wireless Networks*. doi: [10.1007/s11276-010-0271-1](https://doi.org/10.1007/s11276-010-0271-1).
 50. Chowdhury, K., & Felice, M. (2009). Search: A routing protocol for mobile cognitive radio ad-hoc networks. *Computer Communications (Elsevier)*, 32(18), 1983–1997.
 51. Yu, F., Sun, B., Krishnamurthy, V., & Ali, S. (2010). Application layer QoS optimization for multimedia transmission over cognitive radio networks. *Wireless Networks*. doi: [10.1007/s11276-010-0285-8](https://doi.org/10.1007/s11276-010-0285-8).
 52. Zhao, Q., Tong, L., Swami, A., & Chen, Y. (2007). Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: a POMDP framework. *IEEE Journal on Selected Areas in Communications*, 25(3), 589–600.
 53. Clancy, C., Hecker, J., Stuntebeck, E., & O'Shea, T. (2007). Applications of Machine Learning to Cognitive Radio Networks. *Wireless Communications, IEEE*, 4, 47–52.
 54. Li, X., & Zekavat, S. (2009). Traffic pattern prediction based spectrum sharing for cognitive radios. In W. Wang, (Ed.), *Cognitive radio systems* (pp. 77–95). Sciyo.com.
 55. Xiukui, L., & Zekavat, S. (2008). Traffic pattern prediction and performance investigation for cognitive radio systems. In *IEEE Conference on wireless communications and networking, 2008* (pp. 894–899). WCNC 2008.
 56. Hoyhtya, M., Pollin, S., Mammela, A. (2008). Performance improvement with predictive channel selection for cognitive radios. In *First international workshop on cognitive radio and advanced spectrum management, 2008* (pp. 1–5). CogART 2008.
 57. Wang, Z., & Salous, S. (2008). Time series arima model of spectrum occupancy for cognitive radio. In *Seminar on cognitive radio and software defined radios: technologies and techniques, 2008* (pp. 1–4). IET.
 58. Jiang, T., Grace, D., & Liu, Y. (2008). Cognitive radio spectrum sharing schemes with reduced spectrum sensing requirements. In *2008 IET seminar on cognitive radio and software defined radios: Technologies and techniques* (pp. 1–5).
 59. Jiang, T., Grace, D., & Liu, Y. (2008). Performance of cognitive radio reinforcement spectrum sharing using different weighting factors. In *Third international conference on communications and networking in China, 2008* (pp. 1195–1199). ChinaCom 2008.
 60. Kim, H., & Shin, K. G. (2008). Efficient discovery of spectrum opportunities with MAC-layer sensing in cognitive radio networks. *IEEE Transactions on Mobile Computing*, 7(5), 533–545.
 61. Akyildiz, I., & Wang, X. (2005). A survey on wireless mesh networks. *Communications Magazine, IEEE*, 43, S23–S30.
 62. Newman, T., Hasan, S., Depoy, D., Bose, T., & Reed, J. (2010). Designing and deploying a building-wide cognitive radio network test bed. *IEEE Communications Magazine*, 48(9), 106–112.
 63. Sutton, P., Fahmy, S., Nolan, K., Özgül, B., Rondeau, T., Noguera, J., & Doyle, L. (2010). Iris: An architecture for cognitive radio networking testbeds. *IEEE Communications Magazine*, 48(9), 114–12275 (ORBIT Testbed, [Online]. Available: <http://www.orbit-lab.org>. Accessed 2 Dec 2009).
 64. ORBIT Testbed, (Online). (2009). Available: <http://www.orbit-lab.org>. Accessed 2 Dec 2009.
 65. Pawelczak, P., Prasad, R., Xia, L., & Niemegeers, I. (2005). Cognitive radio emergency networks—requirements and design. In *First IEEE international Symposium on new frontiers in dynamic spectrum access networks, 2005* (pp. 601–606). DySPAN 2005.
 66. Hsieh, H., Liu, T., Liao, W., & Ho, T. (2008). Moving toward higher speed WLANs through dynamic spectrum access in the unlicensed band. In *European conference on wireless technology, 2008* (pp. 111–114). EuWiT 2008.
 67. Nan, H., Hyon, T., & Yoo, S. (2007). Distributed coordinated spectrum sharing MAC protocol for cognitive radio. In *2nd IEEE international Symposium on new frontiers in dynamic spectrum access networks* (pp. 240–249). DySPAN 2007.
 68. Kondareddy, Y., & Agrawal, P. (2008). Synchronized MAC protocol for multi-hop cognitive radio networks. In *IEEE international conference on communications, 2008* (pp. 3198–3202). ICC '08.
 69. Zhang, Q., Fitzek, F., & Iversen, V. (2008). Cognitive radio MAC protocol for WLAN. In *IEEE 19th international Symposium on personal, indoor and mobile radio communication, 2008* (pp. 1–6). PIMRC 2008.
 70. Choi, N., Patel, M., & Venkatesan, S. (2006). A Full duplex multi-channel MAC protocol for multi-hop cognitive radio networks. In *1st International conference on cognitive radio oriented wireless networks and communications* (pp. 1–5).

Author Biographies



José Marinho received his B.S. and M.S. degrees in Informatics Engineering from the University of Coimbra, Portugal. Since 1994, he is member of the teaching staff of the Polytechnic Institute of Coimbra (ISEC), being mostly involved with undergraduate courses related to the data networks area. In 2009, he joined the Laboratory of Communications and Telematics (LCT) group of the Centre for Informatics and Systems of the University of Coimbra (CISUC), as a research member. Since 2009, he is a PhD student at the University of Coimbra. His research interests include layer-2 issues in distributed Cognitive Radio scenarios.



Edmundo Monteiro is Associate Professor with Habilitation at the University of Coimbra, Portugal from where he graduated in 1984 and got a PhD in Electrical Engineering—Computer Specialty, in 1995. His research interests are Computer Communications, Service Oriented Infrastructures and Security. He is author of several publications including books and over 200 papers in national and in international refereed books, journals and conferences. He participated in many Portuguese and European research

projects such as RARE, TERENA, COST263, FP5 NoE E-NET, FP6 NoE E-NEXT, COST 290, FP6 IP EuQoS, FP6 SSA OpenNet, FP6 IP WEIRD, FP6 NoE CONTENT, COST TMA, COST WiNeMo, FP7 GINSENG, FP7 MICIE. Besides research projects he has also been involved in consultancy activities with many public and private organizations in Portugal, Europe and African countries. He is member of the Editorial Board of several journals and involved in the organization of many national and international conferences and workshops. He regularly acts as expert for research programs for the European Union, Belgium, Scotland, Switzerland, Russia, Greece, Serbia, Cyprus, Latin America and national organizations in the areas of Research and Innovation, Communication Networks, Information and Communication Technologies, Critical Infrastructures Protection, and Human Resources and Mobility. Edmundo Monteiro is member of IEEE Computer, IEEE Communications, and ACM Communications groups.

ences. He participated in many Portuguese and European research