Cognitive Radio Networks: Not Your Father's Wireless Network

Lance Hester, Ph.D.
Ahmad D. Ridley, Ph.D.

This paper presents the key architectural issues for the creation of cognitive radio networks. The Institute of Electrical and Electronics Engineers (IEEE) describes a cognitive radio as a radio transceiver that intelligently detects whether a particular segment of radio spectrum is currently in use, and whether a radio is able to utilize the temporarily unused spectrum without interfering with the transmissions of other users. Cognitive radios promote open spectrum allocation which is a clear departure from traditional command and control allocation schemes for radio spectrum usage. In short, they allow the formation of "infrastructureless" collaborative network clusters—cognitive radio networks. Research and development of cognitive radio hardware and software, especially at the physical (PHY) layer, is well underway. However, how to transform cognitive radios into functioning cognitive networks is an outstanding area for research exploration. It is with this focus that this paper examines the pitfalls and implementation issues involved in creating cognitive radio networks. It explores dynamic adaptation at layers of the Open System Interconnection protocol stack which includes the PHY, data link, and especially network layers. Emphasis is also placed on describing control and management protocols for dynamic spectrum sensing and possible cognitive radio applications.

Introduction

Whether communicating with others, activating objects like a garage door, or even surfing the Web, people prefer to communicate wirelessly.

> *The Cellular Telecommunications Industry Association (CTIA) stated that one third of all Americans now have cell phones and there are new users signing up at the rate of every 2 seconds—a year-on-year growth rate of 25 percent … . They believe that when handheld computers and smart phones move to faster standards, they will eclipse PCs as the computing platform for the masses … more people will [access] the Internet through these things than any other kind of device … .* [1]

Advances in transceiver technology and integrated circuits have fueled a 10-fold increase in the number of wireless devices used in the home, and a 100x increase in the number of wireless devices used outside the home. The primary drivers of this increase are wireless sensors, pervasive computing, and government and public safety equipment.

This up-tick in wireless usage has meant increased and varied use of radio frequency (RF) across the usable spectrum. This spectrum includes the well-known AM, FM, short-wave and citizens bands, VHF and UHF television channels, as well as hundreds of less familiar bands that serve cellular and cordless telephones, GPS trackers, air traffic control radars, security alarms, radio controlled toys, and the like. [2]

Spectrum utilization

Many of the licensed airwaves are too crowded. Some bands are so overloaded that long waits and interference are the norm. Other bands are used sporadically and are even underused. Figure 1 presents an example of this RF spectrum utilization. Even the Federal Communications Commission (FCC) acknowledges the variability in licensed spectrum usage. According to the FCC, temporal and geographical variations in the utilization of the assigned spectrum range from 15 percent to 85 percent. [3]

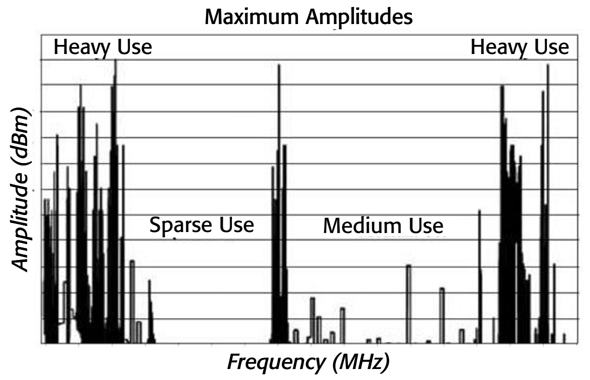

Figure 1. Spectrum utilization [4]

This fluctuating utilization results from the current process of static allocation of spectrum, such as auctions and licensing, which is inefficient, slow, and expensive. This process cannot keep up with the swift pace of technology. In the past, a fixed spectrum assignment policy was more than adequate. However, today such rigid assignments cannot match the dramatic increase in access to limited spectrum for mobile devices. [4] This increase is straining the effectiveness of traditional, licensed spectrum policies.

In fact, even unlicensed spectrum/bands need an overhaul. Congestion resulting from the coexistence of heterogeneous devices operating in these bands is on the rise. Take the license-free industrial, scientific, and medical (ISM) radio band. It is crowded by wireless local area network (WLAN) equipment, Bluetooth® devices, microwave ovens, cordless phones, and other users. Devices participating in unlicensed bands have to do a better job managing user quality of service (QoS).

The limited availability of spectrum and the non-efficient use of existing RF resources necessitate a new communication paradigm to exploit wireless spectrum opportunistically. The new paradigm should support methods to work around spectrum availability traffic jams, make communications far more dependable, and of course reduce interference among users.

The present shortage of radio spectrum can also be blamed in large part on the cost and performance limits of current and legacy hardware. Next generation wireless technology-like software defined radio (SDR) may well hold the key to promoting better spectrum usage from an underlying hardware/physical layer perspective. SDR uses both embedded signal processing algorithms to sift out weak signals and reconfigurable code structures to receive and transmit new radio protocols. However, the system-wide solution is really cognitive radio.

Cognitive radio

Cognitive radio, which was first coined by Mitola in 1999, [5] is a promising approach to achieve open spectrum sharing flexibly and efficiently. Cognitive radio builds on SDR's ability to reconfigure analog output RF and incorporates "self-awareness" and knowledge of transmission protocols, etiquette, and procedures. The result is a cognitive radio able to sense its RF environment and location and then alter its power, frequency, modulation, and other operating parameters so as to dynamically reuse whatever spectrum is available. [4] Cognitive radio nodes adapt their transmission or reception parameters to communicate efficiently without interfering with high priority users or other cognitive radios. In a cognitive radio system, nodes become aware of, learn, and adapt to variations in their interference environment.

From an operator's perspective, cognitive networks maximize the operator's ability to benefit from economies of scale introduced by common hardware platforms and software architectures supporting evolution of radio access solutions, improve time-to-market performance by supporting new service offerings without the need to upgrade the infrastructure, and maximize return-on-investment both in terms of capital expenditures (CAPEX) and operating expenditures (OPEX) by maximizing the exploitation of available/deployed resources. [6]

Cognitive radio research involves each layer of the Open System Interconnection (OSI) protocol stack. A variety of wireless networks have already begun to include cognitive radio aspects. [2] There is an Institute of Electrical and Electronics Engineers (IEEE) standard, IEEE 802.22, based on cognitive radio. [7] IEEE 802.16h [8] is going to bring [cognitive radio] functions into Worldwide Interoperability for Microwave Access (WiMAX) networks for homogenous and heterogeneous network coexistence. [9] Additionally, a number of cognitive radio testbeds have been developed based on different architectures and radio technologies. [9, 10]

The research on cognitive radio covers a wide range of areas, including spectrum analysis, channel estimation, spectrum sharing, and medium access control (MAC). The extreme flexibility of cognitive radios has significant implications for the design of network algorithms and protocols and the applications that will be utilizing them. In fact, they necessitate cross-layer design thinking. Research and development of cognitive radio hardware and software, especially at the physical (PHY) layer, is well underway. However, how to transform cognitive radios into functioning cognitive networks is an outstanding area for research exploration.

In this paper, we examine the pitfalls and implementation issues involved in creating cognitive radio networks. We examine the main features of the various OSI layers of a cognitive radio. Special emphasis is made toward issues related to networking and upper layer protocols—intra-node and inter-node communications, self-forming/self-healing cognitive radio clusters, naming/addressing methodologies, and routing especially to support multi-hop packet relaying among peer radio nodes.

The next section of this paper presents a general overview system-wide perspective of how a cognitive radio node works, including the modified OSI model which includes cross-layer designs. This section is followed by a description of the main aspects of a cognitive radio's physical layer. The key elements associated with a cognitive radio's data link layer are then outlined; the network layer and algorithms for neighbor discovery, node naming, topology formation and management, and routing are introduced; the transport layer and other higher layer implications are described; real-world applications are discussed; and, finally, concluding remarks are presented.

Overview

Non-cognitive wireless devices today employ electronics with the objective to best consume as much of their allocated spectrum as possible. However, in following through on this objective, they are likely to jam (i.e., interfere with the reception of and transmission to) other nearby radios also occupying the same spectrum. [2]

Cognitive radios, on the other hand, are "self-aware" and smart enough to introduce etiquette, sensible operational practices, into RF spectrum operations. Here, self-awareness refers to the radio's ability to learn about itself and its relation to the radio network it inhabits. Cognitive radios intelligently detect and interact with nearby base stations/access points and other spectrum neighbors in such a way that they remain connected using methods that best serve their current needs. It is important to understand that these needs vary with time and situation. Fortunately, these radios have embedded intelligence that enables them to learn from previous interactions with the environment. Based on these interactions, they will adapt their functionality according to different external stimuli. [11]

Cognitive radio users

In a typical cognitive radio scenario, users of a given frequency band are classified into primary users and secondary users. Primary users are licensed users of that frequency band. Secondary users are unlicensed users that opportunistically access the spectrum when no primary users are operating on that frequency band. [9] This scenario exploits the spectrum sensing attributes of cognitive radio. Cognitive radio networks form when secondary users utilize "holes" in licensed spectrum for communications. These spectrum holes are temporally unused sections of licensed spectrum that are free of primary users or partially occupied by low-power interferers. [12] The holes are commonly referred to as white or gray spaces. Figure 2 shows a scenario of primary and secondary users utilizing a frequency band.

In the other cognitive scenario, there are no assigned primary users for unlicensed spectrum. Since there are no license holders, all network entities have the same right to access the spectrum. Multiple cognitive radios co-exist and communicate using the same portion of spectrum. The objective of the cognitive radio in these scenarios is more intelligent and fair spectrum sharing to make open spectrum usage much more efficient.

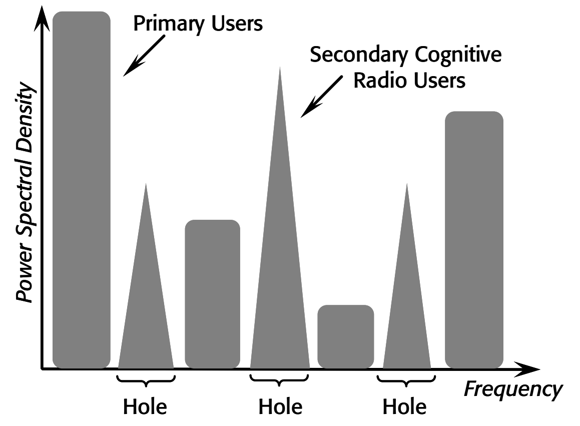

Figure 2. Licensed band scenario showing holes not occupied by primary users—instead, filled by secondary user

Cognitive radio tasks

Basically, a cognitive radio should be able to nimbly jump in and out of free spaces in spectrum bands, avoiding pre-existing users, in order to transmit and receive signals. [2] The main functions of cognitive radios can be summarized as the following four tasks. [4]

- *Spectrum sensing:* Determine which portion of the spectrum is available and detect the presence of licensed users when a user operates in a licensed band.
- *Spectrum management:* Select the best available channel (frequency) for communication.
- *Spectrum sharing:* Coordinate fair spectrum access to this channel with other users.
- *Spectrum mobility:* Vacate the channel when a licensed user is detected while still maintaining seamless communication requirements during the transition to a better piece of spectrum.

These tasks require a large amount of network (and channel) state information that must be shared *simultaneously* between multiple OSI layers of the cognitive radio. Exchanging data among the layers is the case book definition of "cross-layer" design.

Cross-layer design is essential for cognitive radios and especially important for the creation of cognitive networks as will be discussed later. To illustrate, just imagine the case where a cognitive radio has literally only a single antenna for both transmitting and receiving. Under cognitive radio operation, the process of sensing the communication spectrum would have to be stopped so that the radio could switch to a required channel to perform communication, data message exchange. This switch would necessitate cross-layer interaction between the physical layer and upper layers of the cognitive radio's OSI stack.

Another simple example to exemplify the necessity for cross-layer design would be overcoming packet loss on a noisy wireless link. Packet loss can be overcome by local channel coding, as well as end-to-end application coding. Often, the optimal solution requires a combination of both. This optimum can be achieved only if the radio layer, i.e., the PHY and link (MAC) layers, and application layer cooperate to exchange information. [13]

PHY and link layer protocols designed for standard fixed spectrum assigned ad hoc networks will not work. They must be flexible, which means they have to change so that MAC protocols can utilize real-time information from the PHY layer to assign resources for wireless radio nodes. Upper layer assignment policies may be based on various requirements, but spectrum data for making decisions provided by the PHY layer is a priority. [14] Figure 3 [4] depicts the modified OSI stack with cross-layer design ramifications. Note that the figure also illustrates where in the OSI stack the four major cognitive radio functions play a role.

The physical layer

As Figure 3 demonstrates, spectrum sensing is the chief goal of the PHY layer of a cognitive radio. In short, a cognitive radio monitors spectrum bands, captures information, and tries to detect spectrum holes—not an easy feat to accomplish, as will be revealed in this section.

Cognitive radios have to account for situations where potentially you have primary and secondary users occupying the same channel space like in licensed band scenarios or when there are no primaries and every cognitive radio contends with other cognitive and non-cognitive radio for spectrum as in the unlicensed band situation.

This paper concentrates on the situation of primary and secondary users, where the aim is to detect the presence of primary users. The most efficient way to detect spectrum holes is to detect the primary users that are receiving data within communication range of a cognitive radio. By and large, spectrum sensing techniques can be classified as cooperative detection, interference-based detection, and transmitter detection [4] as shown in Figure 4.

Cooperative detection

In cooperative detection, multiple cognitive radios work together to supply information to detect a primary user. This technique exploits the spatial diversity intrinsic to a multi-user network. It can be accomplished in a centralized or distributed fashion. In a centralized manner, each radio reports its spectrum observations to a central controller which processes the information and creates a spectrum occupancy map of the overall network. In a distributed fashion, the cognitive radios exchange spectrum observations among themselves and each individually develop a spectrum occupancy map.

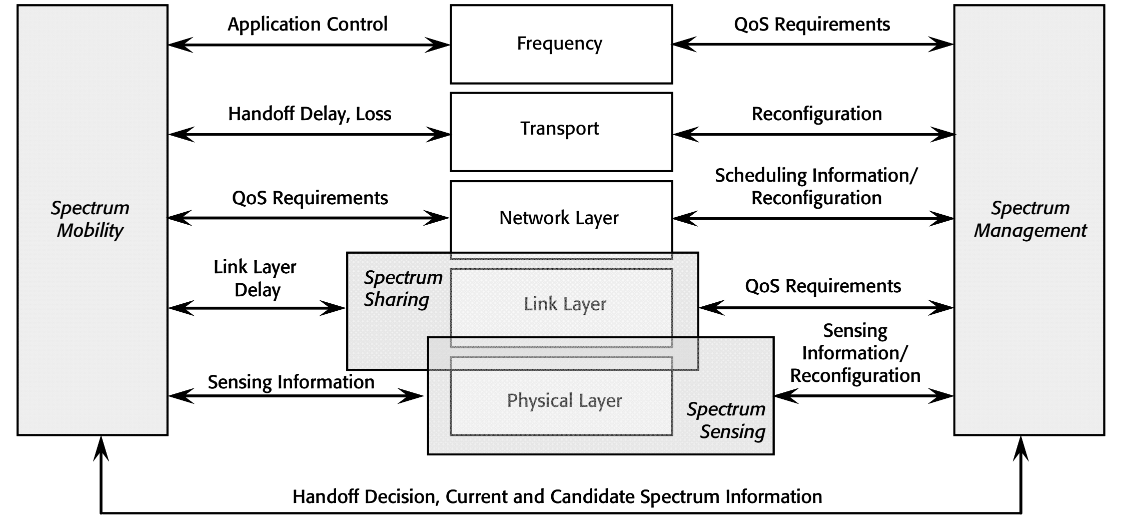

Figure 3. Cross-layer modification of traditional OSI model to better illustrate cognitive radio node inner workings

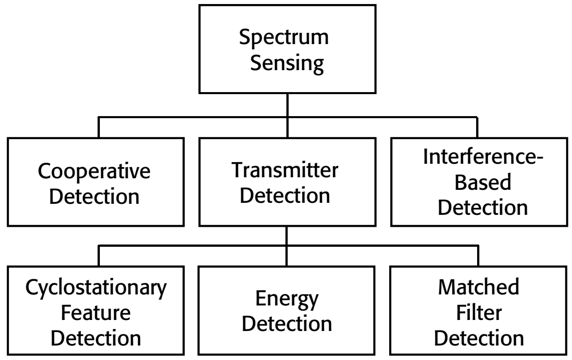

Figure 4. Spectrum sensing techniques

Cooperative detection is advantageous because it helps to mitigate multi-path fading and shadowing RF pathologies which increase the probability of primary user detection. Additionally, it helps to combat the dreaded hidden node problem which often exists in ad hoc wireless networks. The hidden node problem, in this context, occurs when a cognitive radio has good line of sight to a receiving radio, but may not be able to detect a second transmitting radio also in the locality of the receiving radio due to shadowing or because the second transmitter is geographically distanced from it. Cooperation between several cognitive radios alleviates this hidden node problem because the combined local sensing data can make up for individual cognitive radio errors made in determining spectrum occupancy. Sensing information from others results in an optimal global decision.

Shankar [15] points out that cooperative detection requires more network resources because cognitive devices take on the dual role of both data transmission and sensing devices forming essentially a sensor network (for cooperative spectrum sensing) and a data network (operational network). Furthermore, Akyildiz et al. [4] state that the primary receiver uncertainty problem—not knowing the location of the primary receiver—is still unsolved when using cooperative sensing.

Interference-based detection

This method veers from the typical study of interference which is usually transmitter-centric. Typically, a transmitter controls its interference by regulating its output transmission power, its out-of-band emissions, based on its location with respect to other users.

Cognitive interference-based detection concentrates on measuring interference at the receiver. The FCC introduced a new model of measuring interference referred to as interference temperature, as depicted in Figure 5. [16] The model manages interference at the receiver through the interference temperature limit, which is the amount of new interference that the receiver can tolerate. The model accounts for cumulative RF energy from multiple transmissions and sets a maximum cap on their aggregate level. As long as the transmissions of cognitive radio users do not exceed this limit, they can use a particular spectrum band.

The major hurdle with this method is that unless the cognitive user is aware of the precise location of the nearby primary user, interference cannot be measured with this method.

An even bigger problem associated with this method is that it still allows an unlicensed cognitive radio user to deprive a licensee (primary user) access to his licensed spectrum. This situation can occur if a cognitive radio transmits at high power levels while existing primary users of the channel are quite far away from a receiver and are transmitting at a lower power level. Bill Krenik [17] extends the following example to elucidate this point.

> *Take the case of a television station's remote news van. The van sends a signal to the station, which then broadcasts the report to viewers at home. Now, suppose the van is dispatched to an event only a mile or so from the station. The news report will go out of the van at relatively low power across a directional antenna aimed at the station. A [cognitive radio] located outside the directional antenna's influence may not detect the transmission. If the [cognitive radio] decides that the channel is not in use and sends its own signal to a receiver on the far side of the TV station's main transmitter, it can blow the news right off the air. Interference temperature concepts alone cannot effectively protect the licensee in this situation.*

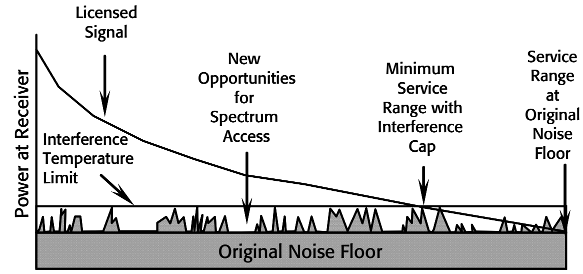

Figure 5. Interference temperature model

A final interference method different from the interference model suggested by Wild and Ramchandran [18] is to detect a primary user by mounting a low-cost sensor node close to a primary user's receiver in order to detect the local oscillator (LO) leakage power emitted by the RF front-end of the primary user's receiver. The local sensor then reports this information to the unlicensed cognitive radio user to help this user develop a spectrum occupancy map.

Transition detection

Here, the cognitive radio attempts to discern areas of used or unused spectrum by determining if a primary user is transmitting in its vicinity. This approach is predicated on detecting not the strongest transmitted signal from a primary user, but the weakest. The idea is that the weakest signal producing primary transmitter would ideally be the one furthest away from the cognitive radio, but still susceptible to RF interference from the radio.

Ghasemi and Sousa [19] describe the basic hypothesis for transmitter detection as:

$$x(t) = \begin{cases} n(t) & H0, \\ hs(t) + n(t) & H1, \end{cases} \tag{1}$$

Here, *x(t)* is the signal received by the cognitive radio, *s(t)* is the transmitted signal of the primary user, *n(t)* is all white Gaussian noise (AWGN) and *h* is the amplitude gain of the channel. *H0* is a null hypothesis, which states that there is no licensed (primary) user signal in a certain band. *H1* is an alternative hypothesis, which indicates that there exists some licensed user signal. [4, 19]

The three main detection techniques which rely on this hypothesis for transmitter detection are described below.

Cyclostationary feature detection. Because modulated signals (i.e., messages being transmitted over RF) are coupled with sine wave carriers, repeating spreading code sequences, or cyclic prefixes all of which have a built-in periodicity, their mean and autocorrelation exhibit periodicity which is characterized as being cyclostationary. Noise, on the other hand, is a wide-sense stationary signal with no correlation. Using a spectral correlation function, it is possible to differentiate noise energy from modulated signal energy and thereby detect if a primary user is present.

Cyclostationary feature detection is a promising option especially in cases where energy detection, described next, is not so effective. However, cyclostationary detection requires a large computational capacity and significantly long observation times. [20]

Energy detection. If a receiver cannot gather sufficient information about the primary user's signal, such as in the case that only the power of random Gaussian noise is known to the receiver, the optimal detector is an energy detector. [4] Energy detection is simple and can be implemented efficiently by using a Fast Fourier Transform (FFT) algorithm. However, there are some drawbacks for energy detection. [20] First, the decision threshold is subject to changing signal-to-noise ratios (SNRs). Second, it can not distinguish interference from a user signal. And third, it is not effective for signals whose signal power has been spread over a wideband.

Matched filter detection. When primary user signal information, such as modulation type, pulse shape, packet format, etc., is known to a cognitive radio, the optimal detector in stationary Gaussian noise is the matched filter since it maximizes the received SNR. The matched filter works by correlating a known signal, or template, with an unknown signal to detect the presence of the template in the unknown signal. Figure 6 [21] provides a graphical representation of this process. Because most wireless network systems have pilots, preambles, synchronization word, or spreading codes, these can be used for coherent (matched filter) detection. [4] A big plus in favor of the matched filter is that it requires less time to achieve a high processing gain due to coherency. The main shortcoming of the matched filter is that it requires *a priori* knowledge of the primary user signal which in a real world situation may not be available.

The data link layer

The job of the cognitive radio's link layer is *spectrum sharing* (as shown earlier in Figure 3). It is apropos to associate the link layer with spectrum sharing because issues related to a radio's access to spectrum are typically concerns of the MAC sublayer. What takes the generic MAC problems of wireless networks to a new level for cognitive radios are obstacles like coexistence between licensed and unlicensed users, dynamic selection of a frequency to transmit upon from a range of available spectrum, and transmitter-receiver handshakes where two or more cognitive radios must agree on a mutual channel upon which to communicate. [4]

Cognitive radio networks provide the opportunity to dynamically change the MAC protocol to suit both the needs of the applications running on the nodes as well as the properties of the environment around them. [10] Choosing the best MAC protocol may be difficult because the best MAC protocol will rarely be best for all nodes involved at any one time. Sometimes, it may be as simple as choosing the best MAC for the greatest number of nodes. While choosing a compatible MAC protocol ensures that a pair of nodes can communicate directly with each other, it is still possible that nodes using incompatible MAC protocols may be able to co-exist in a region in much the same way that existing incompatible wireless protocols co-exist currently (e.g., Bluetooth and WLAN networks). With regards to co-channel interference, MAC protocols should take into consideration that two neighboring channels may co-interfere; so, when channel switching, it might be a good idea to switch to a channel that is a maximal distance in terms of carrier for all neighboring nodes. [14]

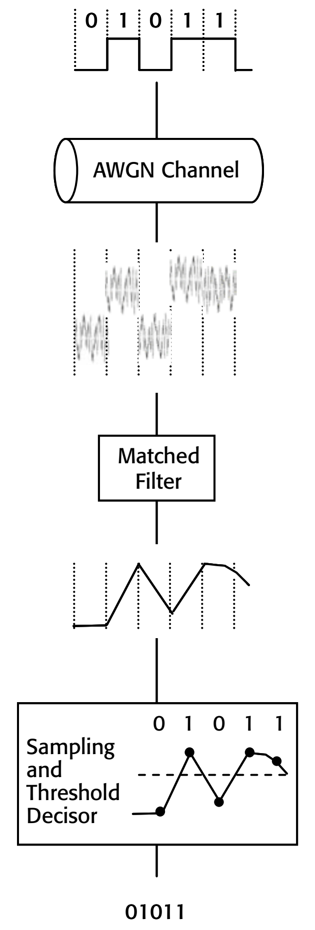

Figure 6. Matched filter

Spectrum sharing policies, or *spectrum etiquette* policies as they are sometimes called, are complicated by the aggregate interference produced by the environment surrounding a cognitive radio. Natural electrical noise (from lightning), electrical power generators, electric motors, radio transmitters, and even automobile ignitions create this interference and their effects change over time. [2] For example, elevators are not very active at midnight, but during the day they are much busier, resulting in the spewing out of electrical noise from their electric motors into the ambient environment.

Methods to mitigate this environmental interference are beyond the scope of this paper. This section is concerned with spectrum sharing methodologies that best attempt to help cognitive radios avoid primary users, and other cognitive and non-cognitive radios. Keep in mind that these methods pay attention to the distinctions between licensed and unlicensed spectrum regimes.

Sharing policies are categorized into three types: architectural, allocation behavior, and access techniques.

Architectural sharing policies

Much research has been done to accomplish spectrum access through centralized and decentralized cognitive radio networks. Authors in [22] and [23] suggest using a centralized spectrum server to dictate spectrum access. In these regimes, radios use a spectrum management protocol to communicate with an impartial centralized spectrum service that obtains information about the interference environment from measurements contributed by multiple cognitive radios. The server offers suggestions for efficient coordination.

Jing and Raychaudhuri [24] advocate a decentralized scheme in which radio nodes do not have any explicit coordination with neighbors. The radios seek equilibrium resource allocation using reactive algorithms that control their transmitted power and data rate. The methods are quite similar to the Transport Control Protocol (TCP) which reactively adjusts source bit-rate over the Internet. Challapali et al. [25] present an alternate decentralized option where cognitive radios use agile wideband radio transceivers to scan a channel and autonomously choose their frequency band and modulation waveform to meet interference minimization criteria without any protocol-level coordination with neighboring radio nodes.

Allocation behavior sharing policies

Nodes may share interference measurements among other nodes, using cooperative spectrum sharing. Or, the nodes can act non-cooperatively using power and rate control techniques.

In many cooperative strategies, nodes look to make use of a reference channel through which they will exchange information about occupied channels. Raychaudhuri and Jing [26] propose a Common Spectrum Coordination Channel (CSCC) approach as a candidate mechanism for implementing spectrum etiquette policies in unlicensed spectrum. Cognitive radios in a cognitive network may use periodic beaconing should they be available to avoid using the same frequencies. Ma et al. [27] propose a dynamic open spectrum sharing MAC (DOSS-MAC) protocol for similar networks in which a common channel is used for signaling.

Zhao et al. [28] point out that in the real world there are a very limited number of global common channels that exist for a network. So, they propose a distributed grouping scheme to solve the common control problem where neighboring cognitive radio nodes may locally share the numerous channels they observe as free with others. The idea is to locally grow spectrum occupancy maps.

This contextual (i.e., time and location) occupancy map construction is useful especially for primary and secondary user situations. If a secondary cognitive radio node knows the exact locations of transmitting primary stations within a licensed band, it can always compare its position with stored coordinates and utilize this information as one of the criterion in the channel/frequency selection process. [14]

Access techniques sharing policies

Overlay spectrum sharing is an interference technique typical of IEEE 802.22 networks. [7] Under this method, secondary cognitive radio users only access the network using portions of spectrum that have not been occupied by a primary licensed user. If a secondary user detects the presence of a primary user on a given frequency band, the secondary user simply vacates that band. This method minimizes interference to primary users by completely avoiding them altogether.

Underlay spectrum sharing gains access to spectrum completely differently. It exploits spread spectrum techniques developed for cellular networks. [29] Once a cognitive radio user acquires a spectrum allocation map (as discussed previously), it transmits messages at a specific power level and at a specific point in the spectrum so that it is regarded as noise by a licensed primary user.

The network layer

Designing cognitive radio network algorithms and protocols is challenging. Cognitive radio networks are not like traditional self-organizing wireless ad hoc networks. They are not designed to work with a single fixed frequency band. They can opportunistically utilize various spectral holes, white spaces, for peer-to-peer communications. Like conventional communication protocols, cognitive network protocols are expected to support a variety of higher layer applications such as voice, data, video, and mobile real-time services, except that they have to compensate for a rapidly changing radio environment, access to multiple radio channels, and PHY and MAC dictated spectrum usage. As previously depicted in Figure 3, creating stable cognitive radio networks involves the distribution and management of inter-node (i.e., PHY-to-link-to-network-layer) and cross-layer information. This fact necessitates that discussions of spectrum mobility and spectrum management be included in this section on network formation and management. The following sections explain the chief functions of the network layer; namely, topology construction, addressing, and routing.

Topology construction

Topology construction involves spectrum detection, neighbor discovery, and topology management—each aspect is discussed below.

Spectrum detection. The first step in constructing a cognitive radio network is the mapping of spectrum occupancy; namely, the “cognitive capability” of the cognitive radio. As discussed earlier, this capability denotes the ability of the radio to capture or sense information about its radio environment beyond just monitoring the power in some frequency band, but also retaining temporal and spatial variations in the radio environment to avoid interference to other users.

Once holes are detected via spectral sensing, they must be characterized, a process coined *spectral analysis*. During spectral analysis, holes are described by their spectrum band information such as operating frequency, bandwidth, and primary user activity. In fact, Akyildiz et al. [4] state that parameters such as interference level, path loss, wireless link errors, link layer delay, and holding time, should be used to represent the quality of a particular piece of spectrum.

Spectral analysis is half of the process. A cognitive radio must also obtain its higher layer user requirements like the bandwidth, data rate, and transmission mode necessary for packet transmission before choosing a spectrum hole upon which to transmit. The combination of user requirements and spectrum characterization leads to the appropriate selection of an operating spectrum band. This process is known as *spectrum decision* and is important for maintaining user QoS.

Neighbor discovery. Generally, radio nodes begin to discover one another via channel scanning or by listening over a control channel should one exist. As discussed earlier, the control channel may be a global control channel produced by a central entity or a local channel created by devices sharing coordinated access to a frequency band accessible to radios in vicinity of one another.

An unassociated node negotiates with existing networked radios for network access and for naming and other network services. In most cases, the unassociated node overhears beacon or pilot messages sent periodically or in an opportunistically beneficial way from networked devices. These beacons relate not only PHY layer information, such as radio parameters for choosing proper operating frequency, transmission power, bandwidth, modulation, data rate, etc., but also network status information. [10]

In a case where a network does not exist, a cognitive radio will initiate the start of its own network usually starting out as a cluster. A cluster is a basic unit of the overall network consisting of two or more cognitive radios. It is a sub-network formed by a group of neighbor nodes sharing common channels and coordinated by a selected node in the cluster called the cluster-head. In this situation, the cognitive radio node starting the cluster is the cluster head.

Cognitive radio cluster formation is somewhat analogous to existing cluster formation algorithms, [30, 31] except that cognitive radio cluster algorithms use more channels than the typical data and possible control channels. These cognitive algorithms adapt to physical topology changes, and they do not necessarily require full topology knowledge. In cognitive situations, a node forming its own cluster will invite adjacent nodes sharing the same channel to join its cluster. [9]

As radios combine to form clusters, the overall radio network is formed from the interconnection of the clusters. Often times, the interconnection of clusters require some cognitive radios that border two separate clusters to act as gateways or bridges. This bridging occurs because many clusters may operate on different frequencies accessible to members of that cluster and not to members of other bordering clusters. Overlapping cognitive radios have access to spectrum in both clusters and, thus, can act as the proverbial "go-between" to link the clusters together.

Topology management. The topologies of clusters are not static—they change, so the global network will have to be reconfigured over time. This topology instability occurs for several reasons. In cognitive radio networks, the available channels for each node fluctuate with regard to the radio environment. Nodes may leave a cluster (spectrum mobility) like in the case that a primary user appears in their vicinity and they must vacate the channel they are using. New nodes may join a cluster. And, whole clusters may merge together as frequencies become available.

The cross-layer spectrum management functions allow for PHY and MAC layer information to be exchanged simultaneously with the network layer so that the network layer will be able to marshal resources from several layers of the OSI stack to assist in network reconfigurability. With this information, network layer protocols can adjust operating parameters "on the fly" for transmissions. Reconfigurability changes are relayed via cross-layer links (reference Figure 3, presented earlier) so that radio operability changes are made without modifications to hardware components. Operating parameters such as operating frequency, modulation, and transmission power can all be changed. For instance, the network layer can determine that for a delay sensitive application, data rate is more important than the error rate. This information can be conveyed via spectrum management functions to the PHY layer so that a modulation scheme that enables higher spectral efficiency would be selected. In addition, cognitive wireless networks will be capable of reconfiguring their infrastructure based on experience (i.e., learned behavior) in order to adapt to continuously changing network environments. [32] They will include topology controls that dictate which nodes should be able to communicate as well as how they should communicate.

A very important area of future research is in alleviating problems caused when radios switch from one channel/frequency to another, commonly known as spectrum handoff. It occurs generally when a secondary user has to vacate spectrum owned by a primary user and the primary user has just shown up, or when channel conditions can no longer sustain communications. The spectrum mobility function ensures that transitions are made smoothly and are transparent to users of the cognitive radio (i.e., a cognitive radio user does not notice any communication degradation while switching from one frequency to another—similar to cell phone handoff). The network protocols should be robust to shift from one mode of operation on one channel to another mode on an alternate channel.

Addressing

Most addressing or network-level addressing schemes use fixed or dynamically derived addresses. The fixed address methods include extensions of the physical MAC address of a cognitive radio or hard-coded address assignment by a central controller device. The dynamic address methods rely on cluster heads to assign node identifiers using an appropriate naming service like that used in IP's Dynamic Host Configuration Protocol (DHCP) service. [10]

Routing

Cognitive radio networks with multi-hop requirements necessitate novel routing algorithms. Generally in ad hoc networks, nodes route packets on each other's behalf, requiring the use of multi-hop routing among nodes to deliver packets. Because available spectrum bands in cognitive radio networks with multi-hop communication are different for each hop, cognitive radios deliver packets from one radio pair on one channel to another radio pair on another channel. Routing between cognitive radios is influenced by control information from a number of sources. [4] Nodes may receive application traffic information from applications running atop them, link capability information from their PHY layers, MAC layer congestion status, and network reachability information from other nodes in their network. [10] Based on all of this information, nodes will need to decide the most appropriate means of delivering the message.

Routing protocols will need to differentiate multi-hop routes based on more than metrics like hop count. In cognitive networks, routing decisions will depend on topology, MAC congestion, and link quality/reliability which will be balanced against complexity and node capability. [10, 14, 32]

Of the three dependencies, link quality/reliability carries even more significance. The notion of "network link" changes because cognitive radios have several adaptation mechanisms like power or bit rate control, and channel adaptation. Will a cognitive radio broadcast at a lower power level or a higher one? Should a radio transmit a message on a single channel or broadcast the message on all channels available to it? These questions are best answered by the application residing on the radio. Fortunately, cross-layer management will help the network layer to specify the parameters for transmitting packets so that the PHY can adapt transmission rate or power.

The combination of waveform-agile radios and the diversity of emerging waveforms allows for radios to use a multitude of radio waveforms and different radio communication protocols (e.g., communicating on a WLAN and then switching to cellular network). Devices will be able to create any kind of communications link they want, with whatever combination of capacity, error-rate, transmission range, etc. The link will no longer be just a "fixed" link, instead, it will be "definable" (i.e., variable) as it will be controlled as necessary by upper layer network protocols. These new links will be tailored to create the type of topology needed by an application and to route data. [13]

High layers

The higher layers of the cognitive radio's OSI stack are responsible for flow and congestion control. Both controls are dictated by the transport layer. The transport layer is affected most by spectrum mobility and MAC performance. Existing transport protocols like the TCP and User Datagram Protocol (UDP) will have to change because they depend on packet loss probability and packet round trip time. Both of these dependencies are influenced by wireless link errors and access technologies which are the purview of the MAC. Cognitive radio's spectrum handoff is a problem because handoff latency can increase packet round trip time which leads to retransmission timeouts. Today's transport protocols might perceive this retransmission timeout as packet loss and invoke congestion avoidance mechanisms resulting in throughput reductions. [4]

Queuing and queue management of packets are also necessary at each cognitive radio because as available spectrum varies over time, a radio may need to store packets while transitioning from one frequency to another during a spectrum handoff. Additionally, during multi-hop communication instances, a separate queue may need to exist for each frequency that a cognitive radio is communicating on for packet relaying.

Applications

Cognitive radio applications will impact existing wireless services and their uses. In the future, cognitive radio technology could transform a cell phone into a WLAN, a laptop into a cell phone, or a cordless phone into a small area wireless access point. [2] Cognitive radio technology has the potential to affect the cost of cellular phone service. A monthly cell phone bill contains charges for items such as leasing radio spectrum, renting cell towers, amortization costs of base site equipment, and interconnection costs among cell sites. [2] Cognitive radio technology will enable cognitive radios to discover, use, and share available radio spectrum optimally, without instructions from a controlling wireless network, thereby, allowing service providers to reduce their leased spectrum and equipment requirements which would trim down overall network costs. This cost decrease or savings could be passed on to consumers in the form of lower monthly service charges.

Another key application space well suited for cognitive radio technology is emergency networks. Cognitive radios are able to use existing unused spectrum, without the need for a dedicated infrastructure, to instantly set up an ad hoc robust network. Ad hoc networks created from cognitive radios could prove extremely useful in disaster assistance situations such as hurricane and tornado relief efforts in coastal and mid-western regions of the United States, respectively. In these scenarios, traditional communication networks are unavailable because of main power outages. Cognitive radios running on batteries or power generators take advantage of the now primary user vacant spectrum to deploy makeshift confined communication regions for public safety operations. Cognitive radio applications will greatly improve emergency response and disaster recovery efforts.

The potential usefulness of cognitive ad hoc networks has not gone unnoticed. In 2004, the FCC made a landmark ruling to promote the creation of ad hoc networks using cognitive radio technology. The FCC recommended that cognitive radio technology be used to create low power ad hoc wireless networks using unused TV bands. [2] Their decision would provide spectrum for cognitive radios across the entire United States, and would, in fact, allow more than 100 MHz of spectrum for cognitive radios in typical urban markets. Cognitive radios, peppered throughout an existing wireless infrastructure, would help capture previously inaccessible spectrum to bolster communications in dense urban areas, while avoiding interference from other radios.

Conclusions

This paper has presented the cross-layer OSI model features of cognitive radios. It described the four primary functions of a cognitive radio: *spectrum sensing*, *spectrum management*, *spectrum sharing*, and *spectrum mobility* and where they participate in the OSI model. Lastly, it explained how these features enable the creation of cognitive radio networks—new types of wireess networks different from today's self-organizing wireless nets. ■

Notes and references

1. Kahney, L., "Smelling Success for Wireless," *Wired Magazine*, November 5, 1999; http://www.wired. com/science/discoveries/news/1999/11/32274.
2. Ashley, S., "Cognitive Radio," *Scientific American Magazine*, February 20, 2006.
3. FCC, ET Docket Number 03-222, Notice of Proposed Rule Making and Order, December 2003.
4. Akyildiz, I., W. Lee, M. Vuran, and S. Mohanty, "NeXt Generation/Dynamic Spectrum Access/Cognitive Radio Wireless Networks: A Survey," *Computer Networks*, vol. 50, no. 13, pp. 2127–2159, 2006.
5. Mitola III, J. and G. Maguire, Jr., "Cognitive Radio: Making Software Radios More Personal," *Personal Communications*, IEEE [see also *IEEE Wireless Communications*], vol. 6, no. 4, pp. 13–18, 1999.
6. Boscovic, D., "Cognitive Networks: A Paradigm for Wireless Communication in Which Networks Adapt, in a Self-aware Manner, Their Topology and Operational Parameters to Fulfill Specific Tasks," Motorola Technology Position Paper, Motorola Corporation; http://www.motorola.com/content.jsp?globalObjectId=6642-9277.
7. Cordeiro, C., K. Challapali, and M. Ghosh, "Cognitive PHY and MAC Layers for Dynamic Spectrum Access and Sharing of TV Bands," Wireless Internet Conference (WICON), 2006.
8. Sydor, J., "Messaging and Spectrum Sharing Between Ad Hoc Cognitive Radio Networks," IEEE International Symposium on Circuits and Systems, 2006.
9. Chen, T., H. Zhang, G. M. Maggio and I. Chlamtac, "Topology Management in CogMesh: A Cluster-based Cognitive Radio Mesh Network," *Proceedings of the IEEE ICC 2007 Workshop (IEEE CogNet 2007 Workshop*), June 24–28, 2007.
10. Raychaudhuri, D., N. B. Mandayam, J. B. Evans, B. J. Ewy, S. Seshan, and P. Steenkiste, "CogNet: An Architectural Foundation for Experimental Cognitive Radio Networks Within the Future Internet," *Proceedings of the ACM/IEEE International Workshop on Mobility in the Evolving Internet Architecture (MobiArch'06)*, December 1, 2006.
11. Demestichas, P., G. Dimitrakopoulos, J. Strassner, and D. Bourse, "Introducing Reconfigurability and Cognitive Networks Concepts in the Wireless World," *IEEE Vehicular Technology Magazine*, vol. 1, no. 1, pp. 32–39, 2006.
12. Haykin, S., "Cognitive Radio; Brain-Powered Wireless Communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 2, pp. 201–220, 2005.
13. NSF-WMPG, "New Architectures and Disruptive Technologies for the Future Internet: The Wireless Mobile and Sensor Network Perspective," *Report of the NSF Wireless Mobile Planning Group (WMPG) Workshop*, New Brunswick, NJ, August 2–3, 2005.
14. Pawelzak, P., "Protocol Requirements for Cognitive Radio Networks," TU Delft, AAF Deliverable WP4.11, June 2005.
15. Shankar, S., "Spectrum Agile Radios: Utilization and Sensing Architecture," *Proceedings of IEEE DySPAN 2005*, Baltimore, MD, November 2005.
16. FCC, ET Docket Number 03-237, Notice of Inquiry and Notice of Proposed Rulemaking, November 2003.
17. Krenik, B., "Clearing Interference for Cognitive Radio," *EE Times*, August 23, 2004; http://www.eetimes.com/showArticle.jhtml?articleID=29100649.
18. Wild, B. and K. Ramchandran, "Detecting Primary Receivers for Cognitive Radio Applications," *Proceedings of IEEE DySPAN 2005*, Baltimore, MD, November 2005.
19. Ghasemi, A. and E. S. Sousa, "Collaborative Spectrum Sensing for Opportunistic Access in Fading Environment," *Proceedings of IEEE DySPAN 2005*, Baltimore, MD, November 2005.
20. Zhang, Q., F. W. Hoeksema, A. B. J. Kokkeler, and G. J. M. Smit, "Towards Cognitive Radio for Emergency Networks," *Mobile Multimedia: Communication Engineering Perspective*, Nova Publishers, 2006.
21. wikipedia.org/wiki/Matched_filter.
22. Raman, C., R. Yates, and N. Mandayam, "Scheduling Variable Rate Links via a Spectrum Server," *Proceedings of IEEE DySPAN 2005*, Baltimore, MD, November 2005.
23. Ileri, O., D. Samardzija, T. Sizer, and N. Mandayam, "Demand Responsive Pricing and Competitive Spectrum Allocation via a Spectrum Policy Server," *Proceedings of IEEE DySPAN 2005*, Baltimore, MD, November 2005.
24. Jing, X. and D. Raychaudhuri, "Spectrum Co-existence of IEEE 802.11b and 802.16a Networks Using the CSCC Etiquette Protocol," *Proceedings of IEEE DySPAN 2005*, Baltimore, MD, November 2005.
25. Challapali, K., S. Mangold, and Z. Zhong, "Spectrum Agile Radio: Detecting Spectrum Opportunities," 6th Annual International Symposium on Advanced Radio Technologies, Boulder, CO, March 2–4, 2004.
26. Raychaudhuri, D. and X. Jing, "A Spectrum Etiquette Protocol for Efficient Coordination of Radio Devices in Unlicensed Bands," *Proceedings of IEEE PIMRC'03*, Beijing, China, September 2003.
27. Ma, L., X. Han, and C. Shen, "Dynamic Open Spectrum Sharing MAC Protocol for Wireless Ad Hoc Networks," *Proceedings of IEEE DySPAN 2005*, Baltimore, MD, November 2005.
28. Zhao, J., H. Zheng, and G. Yang, "Distributed Coordination in Dynamic Spectrum Allocation Networks," *Proceedings of IEEE DySPAN 2005*, Baltimore, MD, November 2005.
29. Huang, J., R. A. Berry, and M. L. Honig, "Spectrum Sharing with Distributed Interference Compensation," *Proceedings of IEEE DySPAN 2005*, Baltimore, MD, November 2005.
30. Lin, C. and M. Gerla, "Adaptive Clustering For Mobile Wireless Networks, Selected Areas in Communications," *IEEE Journal*, vol. 15, no. 7, pp. 1265–1275, 1997.
31. Amis, A., R. Prakash, T. Vuong, and D. Huynh, "Max-Min d-Cluster Formation in Wireless Ad Hoc Networks," *INFOCOM 2000*, IEEE, vol. 1, 2000.
32. Lee, M., D. Marconett, X. Ye, and S. J. B. Yoo, "Cognitive Network Management with Reinforcement Learning for Wireless Mesh Networks," *IPOM 2007*, pp. 168–179, 2007.

About the authors

Lance Hester is an electrical engineer at the Florida Communications Research Lab of Motorola Corporate Labs where his experience includes communication network performance evaluation, network simulation design, and the creation of prototype wireless devices. His research interests include wireless ad hoc communications, wireless sensors networks, and wireless social networking. He received his doctorate degree in electrical engineering and computer engineering from Northwestern University, Evanston, Illinois. Contact him at elh011@motorola.com.

Ahmad D. Ridley is a senior staff engineer at Noblis where his experience includes work on several telecommunications projects under the General Services Administration (GSA) Federal Acquisition Service (FAS) program, including the Washington Interagency Telecommunications Services (WITS) 2001, Federal Technology Service (FTS) 2001, and Networx contracts. His research interests include discrete-event simulation, queuing theory, stochastic optimization, and wireline and wireless data networks modeling. He received his doctorate degree in applied mathematics from the University of Maryland, College Park. Contact him at aridley@noblis.org.